

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ  
СІКОРСЬКОГО»**

Факультет електроніки

(повна назва інституту/факультету)

Акустики та акустoeлектроніки

(повна назва кафедри)

«До захисту допущено»

Завідувач кафедри

Дідковський В.С.

(підпис)

(ініціали, прізвище)

“ ” 20\_\_ р.

**Дипломна робота**

на здобуття ступеня бакалавра

зі спеціальності (спеціалізації) 6.050803 Акустотехніка  
(код та назва спеціальності)

на тему: Оцінки бінаурального сприйняття звуку людино

Виконала: студентка 4 курсу, групи ДГ-51  
(шифр групи)

Деркач Наталія Миколаївна

(прізвище, ім'я, по батькові)

(підпис)

Керівник доцент, Луньова С.А.

(посада, науковий ступінь, вчене звання, прізвище та ініціали)

(підпис)

Консультант \_\_\_\_\_  
(назва розділу) (посада, вчене звання, науковий ступінь, прізвище, ініціали)

(підпис)

Рецензент \_\_\_\_\_  
(посада, науковий ступінь, вчене звання, науковий ступінь, прізвище та ініціали)

(підпис)

Засвідчую, що у цій дипломній роботі  
немає запозичень з праць інших авторів  
без відповідних посилань.

Студент \_\_\_\_\_  
(підпис)

Київ – 2019 року

**Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»**

Інститут/факультет \_\_\_\_\_ електроніки \_\_\_\_\_  
(повна назва)

Кафедра \_\_\_\_\_ акустики та акустoeлектроніки \_\_\_\_\_  
(повна назва)

Рівень вищої освіти – перший (бакалаврський)

Спеціальність (спеціалізація) \_\_\_\_\_ 6.050803 Акустотехніка \_\_\_\_\_  
(код і назва)

ЗАТВЕРДЖУЮ  
Завідувач кафедри  
\_\_\_\_\_ Дідковський В.С.  
(підпис) (ініціали, прізвище)  
«\_\_» \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ  
на дипломний проект (роботу) студенту  
\_\_\_\_\_ Деркач Наталії Миколаївні \_\_\_\_\_  
(прізвище, ім'я, по батькові)**

1. Тема проекту (роботи) Оцінки бінаурального сприйняття звуку людиною  
\_\_\_\_\_

керівник проекту (роботи) \_\_\_\_\_ Луньова С.А., доцент. \_\_\_\_\_,  
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від «\_\_» \_\_\_\_\_ 20\_\_ р. № \_\_\_\_\_

2. Строк подання студентом проекту (роботи) \_\_\_\_\_ 10 червня \_\_\_\_\_

3. **Вихідні дані до проекту** (роботи): \_\_складові артикуляційні таблиці, визначення суб'єктивних та об'єктивних методів оцінки розбірливості мови, макет штучної голови, дані вимірювань бінауральних сигналів, засоби розрахунку коефіцієнтів міжвушної кореляції \_\_\_\_\_

4. **Зміст (дипломної роботи) пояснювальної записки (перелік завдань, які потрібно розробити)** \_\_\_\_\_

1. Аналіз науково-технічної літератури. \_\_\_\_\_

2. Вимірювання коефіцієнтів міжвушної кореляції за допомогою моделі голови. \_\_\_\_\_

3. Порівняння отриманих даних щодо імпульсного та мовного сигналів. \_\_\_\_\_

4. Аналіз характеристик бінаурального сприйняття.

5. Проведення аналізу використаних методів за отриманими результатами.

5. Перелік графічного (ілюстративного) матеріалу (із зазначенням обов'язкових креслеників, плакатів, презентацій тощо) презентація

6. Консультанти розділів проекту (роботи)\*

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв

7. Дата видачі завдання 5 вересня 2018 р

#### Календарний план

№ з/п	Назва етапів виконання дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Збір матеріалів для роботи. Аналіз науково-технічної літератури.	05.09.2018 - 15.10.2018	
2	Вимірювання коефіцієнтів міжвушної кореляції. Аналіз отриманих даних.	16.10.18 - 20.11.18	
3	Теоретичний розрахунок параметрів отриманих даних.	21.11.18 - 10.03.19	
4	Аналіз використаних методів оцінок розбірливості мови.	11.03.19 - 23.05.19	
5	Оформлення пояснювальної записки та презентації.	24.05.2019 - 09.06.2019	

Студент

\_\_\_\_\_  
(підпис)

Н.М. Деркач  
(ініціали, прізвище)

Керівник проекту (роботи)

\_\_\_\_\_  
(підпис)

С.А. Луньова

\* Консультантом не може бути зазначено керівника дипломного проекту (роботи)

## РЕФЕРАТ

Оцінки бінаурального сприйняття звуку людиною// Дипломна робота на здобуття ступеня вищої освіти «бакалавр». Деркач Н.М. Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», факультет електроніки, кафедра акустики та акустoeлектроніки, група

ДГ-51. – К.:НТУУ «КПІ», 2019. с. – 53, рис. – 44, табл. – 6.

Метою роботи є перевірка можливості оцінки розбірливості мови за коефіцієнтами міжвушної кореляції, обрахованими за даними вимірювань бінауральних мовних сигналів, та аналіз отриманих результатів.

В роботі виконаний аналіз бінаурального сприйняття звуку людиною на основі функції міжвушної кореляції. Для вимірювань бінауральної пари сигналів використовується модель штучної голови. Проведений розрахунок коефіцієнтів міжвушної кореляції, отриманих шляхом вимірювань бінауральних мовних сигналів, для оцінки розбірливості мови в приміщеннях. Аналізується можливість оцінки розбірливості мови за коефіцієнтом міжвушної кореляції. Проведено порівняння суб'єктивної та об'єктивізованої оцінки розбірливості мови за допомогою коефіцієнта міжвушної кореляції. Коефіцієнти між вушної кореляції отримані на основі вимірювань бінауральних сигналів в місцях розташування слухачів.

Ключові слова: розбірливість мови, локалізація джерела звуку, методи оцінки розбірливості мови, октані смуги частот, стереозвучання, міжвушна кореляція сигналів.

## ABSTRACT

The assessment of binaural perception of sound by a person// Thesis for a degree of higher education "Bachelor". Derkach N. National Technical University of Ukraine "Kiev Polytechnic Institute named after Igor Sikorsky, Faculty of Electronics, Department of acoustics and acoustoelectronics, group DG-51. - K: NTUU "KPI", 2019. p. - 53, fig. - 44, tab. - 6.

In the work, the analysis and evaluation of binaural perception of sound by a person is carried out. Using the intervertebral correlation function and model of the head. The measurements of intervertebral correlation coefficients, obtained by measuring binaural speech signal, were carried out to assess the locomotive language's fidelity. The possibility of evaluating readability of the language is analyzed by the coefficient of interconnection correlation. A comparison the help of the coefficient of interconnection correlation is carried out of subjective and objectified estimation of readability of the language with. The coefficients are obtained on basis measurements of binaural signals in placement of listeners. Analyzes the possibility of evaluating the intelligibility of speech in terms of the coefficient of inter ear correlation. The idea of using a speech program for objectively assessing the intelligibility of speech in the premises is substantiated. A comparison is made of a subjective and objectivized evaluation of intelligibility of the speech with estimation by means of the coefficient of the correlation. Obtained on basis of measurements binaural signals in the locations of listeners.

The main parameters of language quality are considered to be component and phrasal legibility of the language. The quality of the language in comparison with the quality of the language of the reference path and the quality of speech in real working conditions. Exploring and estimating such parameters with broadband noise is quite difficult, since when using it, we have no phrases or syllables. Consequently, for a more reliable objective assessment of legibility of speech it is expedient to use a speech signal. The purpose of the article is to investigate the possibility of using inter ear correlation obtained by measuring binaural speech signal, in order to assess the readability of the language in the room.

In order the readability of speech in the hall, it possible to use value of the inter ear correlation obtained by measuring binaural signals in listening place.

By comparing the result of articulation test in the room with values of the coefficients between the ear correlation, one can establish scale of IASS values that accord to certain degree of intelligible speech.

Key words: intelligibility of language, localization of sound source, estimation methods, octave of frequency bands, stereo sound inter-sound correlation.

## ЗМІСТ

ПЕРЕЛІК СКОРЧЕНЬ .....	8
ВСТУП .....	9
РОЗДІЛ 1. ВЛАСТИВОСТІ БІНАУРАЛЬНОГО СЛУХУ .....	10
1.1. Аналітичний огляд літератури .....	10
1.2. Локалізація джерел звуку .....	13
1.3. Ефект передування.....	19
1.4. Бінауральні системи звукопередачі .....	22
РОЗДІЛ 2. МЕТОДИ ОЦІНКИ РОЗБІРЛИВОСТІ МОВИ.....	31
2.1. Суб'єктивні методи оцінки розбірливості мови .....	31
2.2. Об'єктивні методи вимірювання рзбірливості мови .....	37
РОЗДІЛ 3. ЛОКАЛІЗАЦІЯ УЯВНОГО ДЖЕРЕЛА ЗВУКУ В ОКТАВНИХ СМУГАХ ЧАСТОТ СЛУХОВОГО ДІАПАЗОНУ .....	38
3.1. Локалізація уявного джерела звуку за взаємною кореляційною функцією стереопари.....	38
3.2. Порівняння ефективності локалізації імпульсного і мовного сигналів в октавних смугах частот.....	40
РОЗДІЛ 4. ОЦІНКА РОЗБІРЛИВОСТІ МОВИ ЗА КОЕФІЦІЄНТОМ МІЖВУШНОЇ КОРЕЛЯЦІЇ МОВНОГО СИГНАЛУ .....	44
ВИСНОВКИ.....	49
СПИСОК ЛІТЕРАТУРИ.....	51
Додаток А.....	53
Спектри мови у різних точках .....	53

## ПЕРЕЛІК СКОРЧЕНЬ

УЗД – уявне джерело звуку

ХН – характеристика напрямленості

АС – акустична система

ПК – персональний комп'ютер

Г1 – гучномовець 1

Г2 – гучномовець 2

L – ліве вухо/канал

R – праве вухо/канал

ITD – interaural time difference

ILD – interaural level differences

УДЗ – уявне джерело звуку

Лб – лівий сигнал бінауральної системи

Пб – правий сигнал бінауральної системи

СС – crosstalk cancellation

HRTF – Head Related Transfer Function



## ВСТУП

Великі зали - надзвичайно складні фізичні системи, і зрозуміти їх в деталях є суттєвою проблемою. Власні коливання, що збуджуються джерелом мають різні показники затухання. Крім цього, час реверберації залежить від типу звукових сигналів, які по різному можуть збуджувати окремі види власних коливань повітряного об'єму приміщень. Через недостатність критерію часу реверберації були запропоновані інші критерії для оцінки акустичних властивостей залів. Питання їх вибору для оцінки акустичних якостей залів на теперішній час є предметом теоретичних та експериментальних досліджень [2].

Можна розробити різні складні фізичні параметри для опису умов в таких просторах, і багато авторів зробили це. Але такі параметри мають не велике практичне значення, якщо вони не належать до суб'єктивної оцінки у тих же самих акустичних умовах. Вкрай важливо проводити контрольовані суб'єктивні тести систематично перевіряти об'єктивні параметри з точки зору суб'єктивних суджень. Через примхи суб'єктивних суджень і фізичну складність великих аудиторій, необхідний великий комплекс досліджень, щоб домогтися істотного прогресу в цій сфері [3].

На сьогоднішній день, є розвиненим використання бінаурального запису для оцінки акустичних якостей приміщень, що імітує сприйняття живої людини. Найчастіше застосовують функцію міжвушної кореляції або іншими словами крос-кореляційну функцію, тобто порівняння сигналів, що надійшли на різні вуха, між собою. Адже основними носіями інформації про звук для людини є час (або часова затримка) амплітуди (або їх різниця) та частота. Після обробки цих даних мозком ми можемо описати, що почули по багатьом критеріям.

У даній роботі пропонується оцінити акустичні якості приміщення за допомогою крос-кореляційної функції, знятої манекеном, що імітує голову людини.

## РОЗДІЛ 1. ВЛАСТИВОСТІ БІНАУРАЛЬНОГО СЛУХУ

### 1.1. Аналітичний огляд літератури

Бінауральні моделі слухового апарату використовуються для моделювання алгоритму локалізації джерела звуку. Моделі класифікують за такими ознаками:

а) за складністю ситуацій:

- локалізація одного джерела в безлуновій камері;
- локалізація джерела в контексті декількох конкуруючих джерел звуку за умов впливу реверберації [3];

б) в залежності від аспектів просторових сцен, які вони досліджують [2]:

- просторовий напрямок і властивості приміщення;
- охоплення слухача;
- відстань до уявного джерела;
- виявлення джерела чи руху слухача.

Також поділяють моделювання, що зосереджується на копіюванні людського сприйняття [4], або на фізіології та модульних моделях периферійної слухової системи [5].

В арифметичних моделях локалізації джерела використовуються сигнали:

- монофонічні;
- бінауральні.

Бінауральні сигнали, зазвичай, надійніші в плані локалізації джерела звуку. Втім вони мінімізують можливість визначення напрямку фронт-тил та кута у вертикальній площині.

Бінауральні сигнали зазвичай характеризують часовою затримкою - ITD та різницею рівнів - ILD.

Модель Джеферса була однією з перших спроб використання ITD [1]. В даній моделі вуха окремо працюють. Для кожного вуха окремо вводиться лінія затримки. Сигнали, що надходять до різних вух, в деякий момент часу перетинаються на ділянці збігу, поширючись в протилежних напрямках. Як наслідок, обидва сигнали потрапляють у зміщену в боковому напрямку ділянку, відносно медіальної площини голови.

На сьогодні, появу УДЗ пояснюється тим, що звукові сигнали, які надходять до різних вух, статистично пов'язані. УДЗ гірше локалізується, якщо зв'язок послаблюється. Поступово, в процесі ослаблення зв'язку, УДЗ розпадається на два незалежних джерела звуку.

В багатьох сучасних моделях локалізації джерела звуку використовують крос-кореляційну функцію (функцію міжвушної кореляції) або ж метод IACC (interaural cross-correlation). Свого часу метод IACC був введений Чері та Сеєрсом [6], як міра схожості двох ідентичних сигналів в залежності від зміщення одного відносно іншого. Така функція відома також як взаємно-кореляційна. Функція широко застосовується в електронній томографії, криптоаналізі та нейрофізіології. В задачах локалізації визначається за результатами вимірювань імпульсних характеристик приміщення на лівому і правому вусі, в різних точках приміщення. Оцінюється як кореляційна функція нормована на взаємну енергію відповідних сигналів:

$$IACF(\tau) = \frac{\int_{-\infty}^{\infty} p_l(t) p_r(t + \tau) dt}{\sqrt{\int_{-\infty}^{\infty} p_l^2(t) dt \cdot \int_{-\infty}^{\infty} p_r^2(t) dt}}, \quad (1)$$

де  $\tau$  – часова затримка,  $p_l(t)$  та  $p_r(t)$  – звуковий тиск на лівому та правому вухах.

Використання бінаурального запису є досить поширене явище в наш час. Його використовують для оцінки можливості локалізації джерела звуку людським організмом, а саме, визначення просторових параметрів області, в якій присутній ефект сумарної локалізації.

Час запізнення сигналів, які надходять на різні вуха складає  $\pm 1\text{мс}$ . Тому і розрахунки, традиційно, проводяться, виходячи з цього припущення [2], [11], [12].

По-перше час затримки відповідає знаходженню джерела звуку в боковому напрямі слухової системи. По-друге даний підхід виник під впливом «ефекту передування» або ефекту Хааса, який відтворює один з існуючих механізмів локалізації звуку в приміщенні [10]. Це явище зводиться до наступного: якщо з різних напрямків надходять звуки з короткими інтервалами затримки, то за правилом “першої хвилі” відбувається локалізація, тобто за першим звуком, що надійшов. Якщо інтервал між двома короткими звуками стає меншим за 1 мс, то ефект не виявляється і відбувається “сумарна локалізація”. Інтервал сумарної локалізації знаходиться в межах 5 мс для коротких імпульсних сигналів. Для мовних та музичних сигналів часова затримка для сумарної локалізації ще більший, і за дослідженнями різних авторів складає від 40 мс до 60 мс [2], [11], [12]. Крім того, за даними досліджень, виконаних Ю. А. Ковалгінім [11], при часі затримки  $0,8 \div 1,2\text{мс}$  переміщення УДЗ залежить від типу сигналу.

Амплітуда коливання місцезнаходження УДЗ стає незначною тільки при  $\tau = 7\text{мс}$ . При такому часі затримки УДЗ практично досягає положення крайнього джерела звуку, незалежно від типу сигналу. Подальша поведінка УДЗ залишається невизначеною, але остаточне розпадання його на два джерела звуку настає при значно більшому часі затримки. Такі дані отримані на основі суб’єктивних оцінок експериментальних випробувань для приміщень малого об’єму і малих розмірів бази стереосистеми (до 2,8 м) [11].

З огляду отриманих результатів досліджень, виникають питання:

- як залежить можливість локалізації джерела звуку при відтворенні більш складних сигналів - мовних і музичних;
- за якими саме об’єктивними параметрами можливо оцінювати зону існування УДЗ.

Бінауральним слухом називається його здатність визначати напрям приходу звукової хвилі, тобто локалізувати положення джерела звуку в просторі. Ця здатність досягається завдяки просторовій несуміщеності двох вух в поєднанні з екрануючим впливом голови. Тому завжди має місце не ідентичність збудження правого і лівого вуха. Цей факт забезпечує людині можливість сприймати просторовий звуковий світ і оцінювати переміщення джерел звуку в цьому просторі.

До числа основних властивостей бінаурального слуху можна віднести:

- просторову локалізацію,
- ефект передування,
- бінауральне підсумовування гучності,
- бінауральне демаскування.

Такі можливості слуху досягаються завдяки трьом чинникам:

1. Часовий - виникають через розбіжність моментів впливу однакових фаз звуку на ліве і праве вухо.

2. Амплітудний - виникають через неоднакову величину звукових тисків на ліве і праве вухо внаслідок дифракції звукової хвилі навколо голови.

3. Спектральний - виникають через різницю в спектральному складі звуків, сприймаються лівим і правим вухом, внаслідок неоднакового екрануючого впливу голови і вушних раковин на НЧ і ВЧ складові спектра складного звуку.

## 1.2. Локалізація джерел звуку

**Азимутна локалізація джерел звуку.** Якщо під деяким кутом  $\varphi$  до медіанної площини голови 1-1 слухача знаходиться джерело звуку (рис.1.1), то фронтальна локалізація включає визначення азимута (кута  $\varphi$ ) і відстані  $l$  до джерела звуку. Внаслідок дифракції звукової хвилі навколо голови

слухача і частотно-залежного загасання цієї хвилі з відстанню звуку до правого і лівого вух слухача приходять не однаковими. Вони відрізняються за рівнем інтенсивності на величину  $\Delta N$  і за часом на величину  $\Delta t$ , а також є функцією азимута і частоти  $\varphi$ . Ці параметри і є носіями інформації про локалізацію джерела звуку.

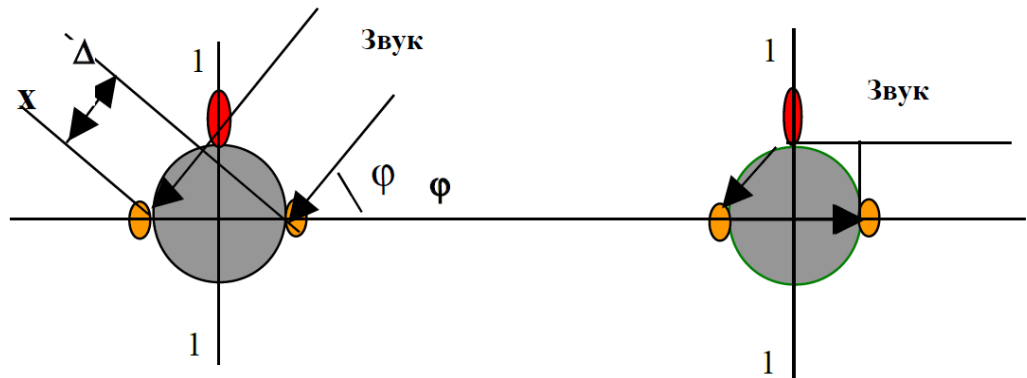


Рис 1.1 Різниця приходу звукової хвилі у часі

Різниця часу  $\Delta t$  приходу однакових фаз звуку до вух визначається рівністю  $\Delta t = \frac{\Delta x}{C_0}$  де  $\Delta x$  - різниця ходу звукової хвилі до лівого і правого

вуха,  $C_0$  – швидкість поширення звуку в повітрі (340 м / с). У свою чергу  $\Delta x$  залежить від розмірів голови людини, точніше від відстані між вухами  $B$ , яке

в середньому дорівнює 18 см: 
$$\Delta x = \frac{B}{2} + \frac{B \cdot \pi}{4}$$

Максимальна різниця ходу досягає 23 см при розташуванні джерела під кутом 90 градусів (збоку). В цьому випадку вона більше середньої відстані між вухами  $B = 18$  см і менше найліпшої відстані навколо голови, що дорівнює 28 см (Рис. 1.1). При такій різниці ходу максимальний час затримки дорівнює 0,63 мс. Однак, це справедливо тільки для синусоїдальних звуків з частотою не вище 800 Гц. На високих частотах час затримки стає більше половини періоду коливання, тому поняття фази втрачає сенс. Низькі частоти мають довжину хвилі більше, ніж діаметр голови слухача, тому вони оминають голову і не дають акустичної тіні. Звуки високої частоти мають

довжину хвилі менше ніж діаметр голови слухача, тому вони не проходять до лівого вуха. Виникає, внаслідок явища дифракції, акустична тінь.

З досвіду випливає, що чисті тони високих частот (понад 8 кГц) майже не піддаються локалізації. Також слабо виражена здатність локалізації на частотах нижче 300 Гц. На частотах менше 150 Гц локалізація відсутня взагалі. На цій підставі не має значення, де розміщувати НЧ акустичні системи при стереозвучанні.

**Найбільша точність локалізації** досягається при сприйнятті складних звуків і звукових імпульсів. При цьому важливим фактором є спектральний склад звуків. Так, якщо звук містить НЧ і ВЧ складові і діє під кутом 90 градусів на праве вухо, то в лівому вусі ВЧ складових не буде, у зв'язку з тінню.

**Розрізнявальна здатність зміни положення** джерела звукових імпульсів дорівнює  $3^0$ . Цю величину слід вважати бінауральною роздільною здатністю слуху для фронтального напрямку. **Точність локалізації** джерел звуку, розташованих зліва і справа, значно менше і становить приблизно  $12^0$ . Для тилового напрямку ця величина дорівнює близько  $6^0$ .

При локалізації джерел звуку параметри  $\Delta N$  і  $\Delta t$  повністю взаємозамінні. При оцінці азимута орган слуху обмінює еквівалентне значення інтенсивностной різниці  $\Delta N_{\text{эк}}$  на часову різницю:

$$\Delta N_{\text{эк}} = K_{\text{э}} \Delta t ,$$

де  $K_{\text{э}}$ - коефіцієнт еквівалентності. Сумарне значення еквівалентної інтенсивності  $\Delta N_{\text{СУМ}}$  визначається сумою

$$\Delta N_{\text{СУМ}} = \Delta N + K_{\text{э}} \Delta t$$

Вона є додатковим фактором для визначення азимута.

Величини  $\Delta t$  і  $\Delta N$  є не тільки лінійними функціями азимутального кута, але залежать також від частоти. Вони змінюються при переході від однієї частотної групи слуху до іншої, залишаючись приблизно постійними в межах однієї частотної групи. На низьких частотах (нижче 500 Гц)  $\Delta N \ll$

$K, \Delta t$ , тобто велику роль відіграє часовий фактор. У діапазоні середніх частот від 500 до 5000 Гц обидва ці чинники приблизно в рівній мірі сприяють створенню відчуття напрямку.

Таким чином, при оцінці азимута джерела звуку голова і вушні раковини виконують функцію просторового фільтра. Судження про величину кутового зміщення джерела звуку від медіанної площини пов'язане з оцінкою слуховою системою часових і інтенсивносних відмінностей пари бінауральних звуків. Така оцінка здійснюється в кожній частотній групі слуху і порівнюється з завченими, набутими в результаті досвіду зразками, які зберігаються в слуховій пам'яті. Все це і дозволяє визначити азимут.

**Локалізація джерел звуку в вертикальній площині.** Здатність визначати напрямок приходу звуку в вертикальній площині у людини розвинена дуже слабо і складає всього 10-15 градусів. Ця здатність пов'язана з орієнтацією і формою вух.



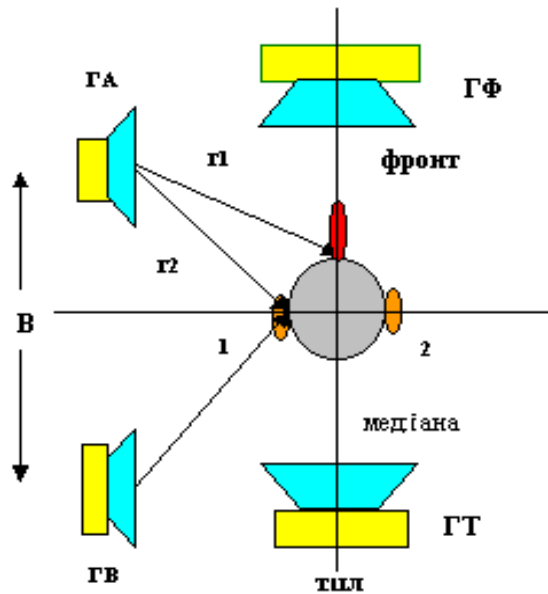


Рис 1.2. Частотна залежність різниці рівней звукового тиску у лівого і правого вуха за умов нормаління падіння

**Механізм сприйняття звуку в площині фронт-тил.** Коли одне джерело випромінювання (ГА або GB) розташоване у фронтальній площині або тилівій з азимутом, тоді відстані поширення звуку до вух різні ( $r_1$  і  $r_2$ ). У цій ситуації в локалізації приймають участь обидва параметри:  $\Delta t$  і  $\Delta N$ . Якщо джерело звуку розташований в медіанній площині (ГФ або GT), тоді значення  $\Delta t$  і  $\Delta N$  для них практично однакові ( $r_1 = r_2$ ) і неможливо оцінити азимутальний кут  $\varphi$  джерела звуку щодо медіанної площини (рис. 1.2). Незважаючи на це, локалізація виявляється безпомилковою завдяки додатковому спектральному аналізу бінауральної пари звуків.

Орган слуху людини має ніби два різних механізми для оцінки місця розташування джерела звуку в просторі. Один з них (фронт-тил) визначає, чи знаходиться джерело спереду або ззаду слухача, другий визначає напрямок на джерело звуку щодо медіанної площини (кут  $\varphi$ ). Якщо джерела звуку знаходяться в медіанній площині, то  $r_1 = r_2$ .

Ключовим моментом для розуміння роботи механізму фронт-тил при цьому є залежність, показана на рис. 1.3. Ця залежність є графіком зміни

різниці рівнів звукового тиску  $\Delta N_{\phi T}$  на праве і ліве вухо в залежності від частоти між рівнями тисків фронтального  $\Delta N_{\phi}$  і тилового  $\Delta N_T$  гучномовців

$$\Delta N_{\phi T} = N_{\phi} - N_T$$

Як видно з графіка, на одних частотних смугах  $N_{\phi} > N_T$ , на інших навпаки.

Для впевненої фіксації слухачем фронтального або тилового напрямків досить мати відмінність гучномовців за звуковим тиском більше 1,5 - 2 дБ. Умови роботи фронт-тил погіршуються, якщо гучномовці знаходяться *поза* медіанної площини. Цей механізм взагалі не працює при  $\phi = 90$  і  $270$  градусів.

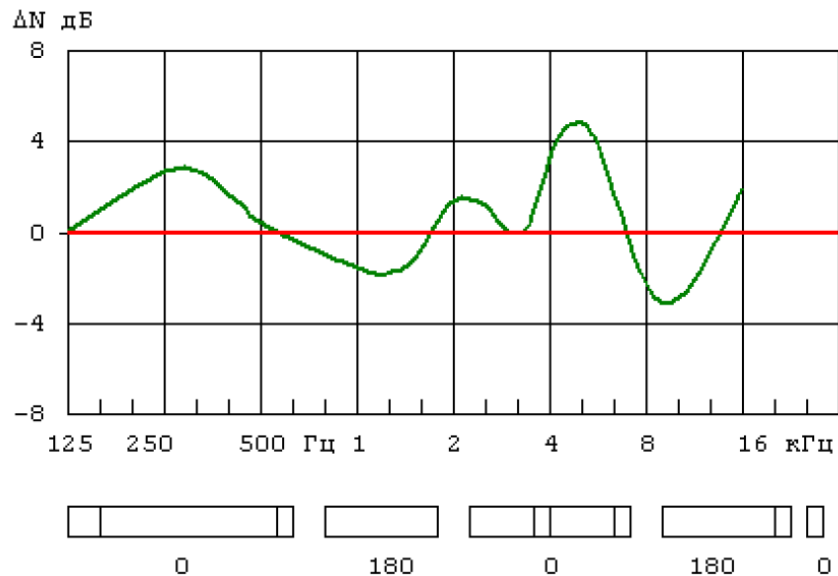


Рис. 1.3 Сприйняття звуку в медіанній площині

**Локалізація джерел звуку по глибині.** Поряд з визначенням напрямку на джерело звуку слухач впевнено оцінює відстань до джерела звуку.

1. На середніх відстанях до джерела звуку 3-15 м наближення і віддалення джерела звуку супроводжується зміною його інтенсивності. У вільному звуковому полі на НЧ збільшення відстані вдвічі супроводжується зменшенням рівня звукового тиску на 6 дБ. На слух відстань завжди визначається менше, ніж вона є насправді. Помилка зростає зі збільшенням відстані.

2. На близьких відстанях менше 3 м на глибинну локалізацію впливає дифракція на вушній раковині і голові, тобто позначається різниця рівнів інтенсивності і часові затримки. Точність глибинної локалізації при зміщенні широкосмугового джерела від 50 до 150 см не перевищує 15-30%.

3. На великих відстанях понад 15 м починають впливати загасання, залежні від відстані, яку проходить звукова хвиля. При цьому ВЧ складові загасають швидше, змінюється тембр. На поширення звуку впливає вологість повітря і напрямок вітру. Збільшення амплітуди НЧ складових спектра звуку викликає відчуття наближення джерела звуку. Штучне зменшення ВЧ складових спектра сприймається як віддалення джерела звуку.

4. На відміну від азимутальної локалізації глибинна локалізація можлива при моноуральному прослуховуванні, однак, бінауральне сприйняття істотно підвищує точність оцінки відстані. При цьому орган слуху підсвідомо оцінює величини  $\Delta t$  і  $\Delta N_{cp} / \Delta N$ , де  $N_{cp}$  - середнє значення інтенсивностей звуків, діючих на ліве і праве вухо. При відстанях більше 10 м ці фактори не працюють.

### 1.3. Ефект передування

У приміщеннях, де поряд з прямим звуком на слухача впливає значне число відбитих хвиль, важливим фактором, що стимулює глибинну локалізацію, є ефект передування. Суть цього ефекту полягає у відокремленні слуховою системою прямого звуку від його ревербераційних продовжень. Судження про направлення на джерело звуку формують прямі звуки, в той час як частина наступних за ними повторень на інтервалі від 1-5 до 30-50 мс пригнічується слуховою системою. При великих затримках придушення не відбувається. Завдяки ефекту передування виявляється можливою слухова пеленгація джерела звуку в приміщеннях, де в точку

розташування слухача поряд з прямим звуком надходить значна кількість хвиль, відбитих від його поверхонь.

**Бінауральне маскування і демаскування.** Коли звук, що маскується і маскувальний звук надходять в різні вуха, то слід говорити про бінауральне маскування. Величина зсуву слухового порога при бінауральному маскуванні завжди набагато менше, ніж при моноуральному маскуванні, і проявляється вона в більшій мірі на високих частотах.

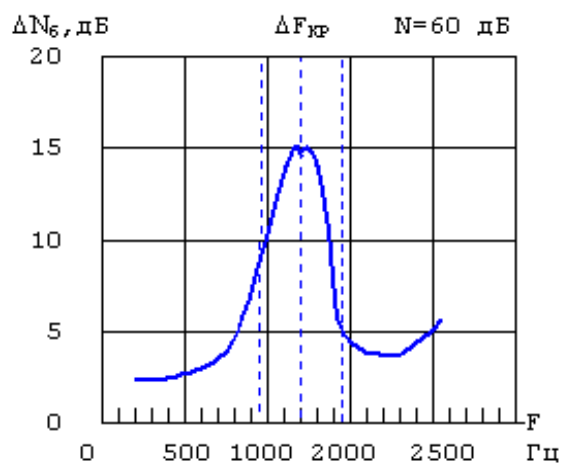


Рис 1.4. Графік бінаурального маскування

Бінауральне маскування виникає тільки за умови, якщо час впливу маскувального звуку не менше 250 мс. Найбільш виразне маскування коли звуки близькі по частоті. При цьому на відміну від моноурального маскування відсутні биття і маскування виражене в дуже вузькій смузі частот, цей діапазон збігається з шириною критичних смуг слуху (рис. 1.4). Ступінь бінаурального маскування наростає, коли обидва звуки мають пульсуючий характер, що характерно для мови і музики.

Здатність налаштовуватися на один з багатьох одночасно діючих джерел звуків становить одну з дивовижних особливостей людського слуху. Однією з найважливіших складових цієї особливості слуху людини є

бінауральне демаскування звуків. Бінауральним демаскуванням називають зниження порогу маскування при відокремленні звуків із сукупності діючих звуків. Це досягається шляхом додаткової обробки вхідної інформації як в периферійних, так і центральному відділах слухової системи і мозку. Зазвичай звуки мають спектри, які перекриваються, тому класична теорія фільтрів з її областями пропускання і загасання тут абсолютно не діє. Вуха є значно більш тонким інструментом. Бінауральне демаскування передбачає використання слуховою системою тимчасових пар бінауральних звуків, які відповідають окремим джерелам для їх поділу один від одного. Деякі дослідники вважають, що в основі цього явища лежить чутливість до зсуву фаз між звуками на частотах нижче 1500 Гц. Якщо в стереотелефони подати одночасно один і той же шум, і один і той же маскувальний звук, то крім шуму нічого не почуємо. Якщо звук, який маскується, подати тільки в одне вухо, то ми його будемо чути, його рівень суб'єктивно підвищується. Цей ефект і називається бінауральним демаскуванням. При цьому шум і звук локалізується в різних місцях голови. Шум в середині голови, а звук ближче до вуха, на яке він подається. Таким чином, демаскування здійснюється через різну суб'єктивну локалізацію шуму і звуку. Аналогічне явище відбувається при тихій розмові в галасливій аудиторії, наприклад, на дискотечі. Людина мимоволі повертає голову і знаходить положення, при якому в обидва вуха надходить однаковий шум, який сприймається в середині голови. Звук розмови локалізується в іншому місці ближче до вуха і він добре чутний. Цей механізм працює тільки при наявності в спектрі звуку НЧ складових. Коли звук, який маскують і маскувальний звук надходять на різні вуха, цей процес називають бінауральним маскуванням. Величина зсуву слухового порогу при бінауральному маскуванні завжди набагато менше, ніж при моноуральному маскуванні і проявляється вона в більшій мірі на високих частотах. Бінауральне маскування виникає тільки за умови, якщо час впливу маскувального звуку не менше 250 мс. Найбільш виражено маскування коли звук, який маскують і маскувальний звук близькі по частоті. При цьому на

відміну від моноурального маскування відсутні биття і маскування виражено в дуже вузькій смузі частот, цей діапазон збігається з шириною критичних смуг слуху.

#### 1.4. Бінауральні системи звукопередачі

Ідея бінауральної передачі представлена на рисунку 1.5. Якщо помістити два мікрофони в вушні канали акустичного манекена і після цього передати сигнали, без якихось змін підвести за допомогою головних телефонів до відповідних вушних каналів слухача. В такому випадку слухач знаходиться ніби в початковому звуковому полі. При цьому, якщо манекен і слухач мають однакові розміри голів, а також схожі за формою та акустичними характеристиками, то і сигнали, сприйняті мікрофонами

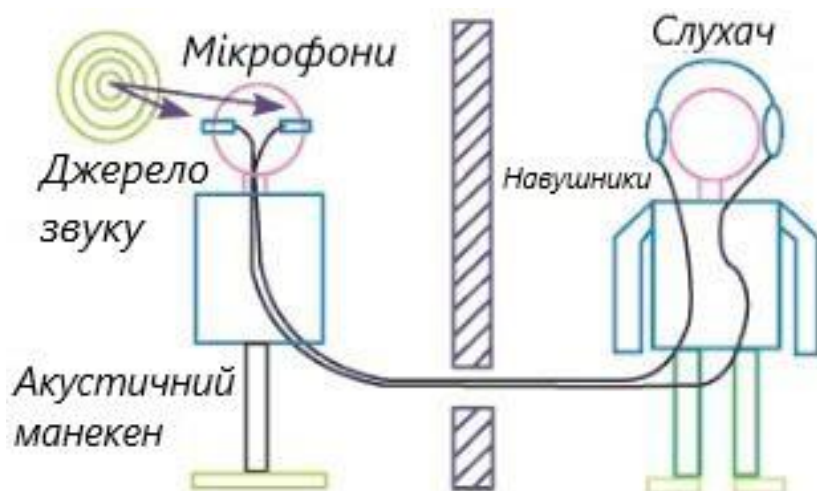


Рис 1.5. Система для бінауральної передачі

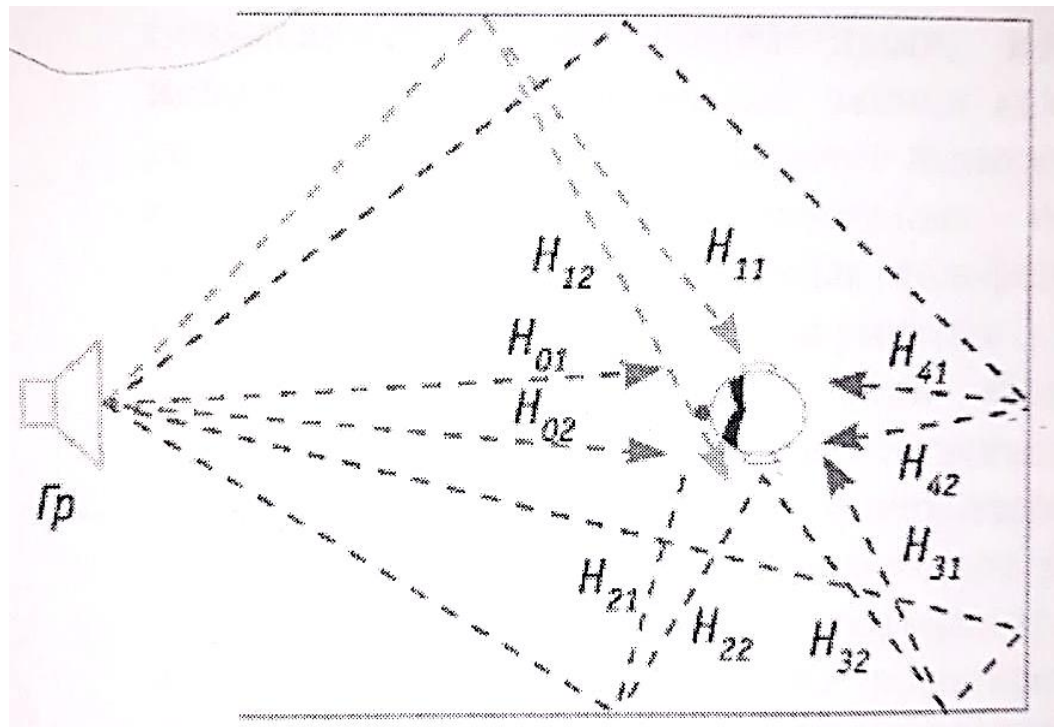
манекенами і вухами слухача, будуть ідентичні.

Системи звукопередачі, які реалізують даний принцип, називаються бінауральними. Вони мають здатність створювати тривимірну, винесену за межі голови слухача просторову звукову панораму, яка створює повну ілюзію присутності слухачів концертному залі, а це означає, що існує здатність забезпечувати вищу, ніж при стереофонії, якість звучання.

В бінауральній звуковій системі формується лівий ( $L_6$ ) і правий ( $P_6$ ) сигнали, тотожні тим, які впливали б на входи слухового аналізатора слухача при його розташуванні в первинному звуковому полі, наприклад в концертному залі. Новими елементами в порівнянні зі звичайною стереофонічною системою, яка в своєму складі має пульт звукорежисера (ПЗ) та стереофонічний підсилювач (СП), тут є бінауральний регулятор напрямку (БРН) на передавальній стороні і біофонічний процесор (БП) на стороні відтворення звукових сигналів, якщо відтворення сигналів  $L_6$  та  $P_6$  виконується за допомогою гучномовців  $G_1$  та  $G_2$ .

При використанні гучномовців для відтворення бінауральної пари сигналів  $L_6$  та  $P_6$  необхідний додатковий пристрій, який називається біофонічним процесором. Він забезпечує, при гучномовному відтворенні, підведення до лівого та правого входів слухового аналізатора бінауральної пари сигналів  $L_6$  та  $P_6$  в своїй первинній, неспотвореній формі.

Сукупність відбитих сигналів (рис.6) створює ревербераційний процес житлового приміщення, вона впливає на ліве та праве вухо слухача, змінюючи пару сигналів  $L_6$  та  $P_6$ . В якості БРН часто використовують «штучну голову», в вушних проходах якої, на місці барабанної перетинки встановлені приймачі звуку – мікрофони.



**Рис 6.** Схема відбитих від поверхонь сигналів

**Недоліки та переваги бінауральної системи.** Основні переваги та недоліки бінауральної системи передачі звуку можна сформулювати таким чином:

1. У разі компенсації заважаючих факторів можлива чітка локалізація уявних джерел звуку (УДЗ) в межах своєї горизонтальної площини: спереду, зліва, справа, ззаду. Та все ж УДЗ, розташовані спереду, здаються більш зосередженими, а в тилівій частині – більш розмитими. За умов використання бінаурального запису УДЗ, у разі застосування головних телефонів, виявляються винесеними за межі голови слухача в те місце, де розташовані дійсні джерела звуку. На відміну від цього, під час прослуховування через головні телефони сигналів звичайної стереофонії УДЗ відчуються як розташовані всередині голови слухача. При бінауральній стереофонії оцінка відстані до джерела звуку, як і за умов природнього прослуховування, також не є точною.

2. Наявність реверберації, як в первинному, так і вторинному приміщеннях, ускладнює оцінку азимута УЗД в тилівій області, де слухачі замість істинного напрямку часто вказують відповідний йому дзеркальний



фронтальний напрямок. Дане явище виникає особливо часто, коли час стандартної реверберації в приміщенні перевищує 0.3 с. Щодо відчуттів акустичних властивостей первинного приміщення, пов'язаних з передачею глибинної локалізації (гулкості, теплоти, м'якості), то вони є правдоподібними. Бінауральній системі притаманне точне відновлення просторової структури первинного ревербераційного процесу в порівнянні зі звичайною стереофонією.

3. Правильна передача просторової інформації за умов використання гучномовців можлива тільки для однієї точки приміщення. За допомогою біофонічного процесора скомпенсована дія факторів, які виступають завадами. Навіть незначне зміщення (9-15 см) центра голови слухача вліво або вправо від цієї точки призводить того, що локалізація УЗД є неможливою поза фокусу гучномовців.

Не викликає сумніву той факт, що в оптимальній точці прослуховування бінауральна система забезпечує звучання, створює ефект присутності слухача в концертному залі з хорошими акустичними властивостями. Однак область її застосування є досить обмеженою: відтворення за допомогою головних телефонів, переносна апаратура радіомовлення і звуковідтворення, комп'ютерне моделювання. Бінауральна звукова система не є вдалим варіантом за умов колективного прослуховування.

Перша вдала спроба реалізації бінауральної системи з гучномовним відтворенням є система TRADIS, створена Б.Бауером і реалізована П.Дамаске та В.Меллертом. Вхідними сигналами є бінауральна пара  $L_6$  та  $P_6$ . Тут здійснюється компенсація тільки перехресних сигналів, один з яких надходить від лівого гучномовця  $\Gamma_1$  і сприймається правим вухом, а інший – від правого гучномовця  $\Gamma_2$  і сприймається лівим вухом. Саме наявність цих сигналів робить неможливим локалізацію УЗД поза бази гучномовців.

До слухача, який знаходиться в житловому приміщенні, окрім сигналів прямих звуків, надходять також звукові хвилі, відбиті від поверхонь

приміщення. Сигнали додаються, спотворюють вхідні сигнали, роблячи їх нетотожними і порушуючи тим самим принцип бінауральної звукопередачі. До того ж першими за часом надходження і найвищими за рівнем є ті відлуння, які приходять з тієї частини середовища, де розташовані гучномовці. Ці відлуння ускладнюють сприйняття тилових пеленгів, стимулюючи локалізацію дзеркального напрямку. Дане явище стає непомітним, якщо час перевищує 0.3 с, тобто коли існує чітко виражена дискретна частина ревербераційного процесу. На завершальній ділянці ревербераційного процесу в кожен момент часу приходить значна кількість відлунь з різних напрямків. Рівні відлунь малі, тож вплив цієї ділянки на слуховий пеленг не має бути значним.

Поява дзеркальних пеленгів пов'язана з впливом саме перших відлунь, які надходять завжди з тієї області простору, де розташовані гучномовці. Для їх компенсації необхідні спеціальні електронні схеми, досить складні за своєю структурою.

Наступними вагомими недоліками бінауральної системи: правильна передача просторової інформації є можливою тільки для однієї точки прослуховування. Навіть незначні зміщення (10-15 см) центра голови слухача вліво і вправо від цієї точки приводить до того, що слухова пеленгація УЗД є неможливою. На думку японських дослідників, застосування бінауральних регуляторів напрямку допускає поперечне зміщення слухача не більш, ніж на 30 см.

За наявності реверберації у вторинному приміщенні найкращі результати бінауральна система звукопередачі з двома гучномовцями забезпечує, навіть за умов малих розмірів бази і досить близькому розташуванні до нього слухачів.

Амплітудно-частотні характеристики лівого (правого) і перехресних каналів процесора: величина  $t = 200 \text{ мкс}$ ; відстань гучномовців до найближчої стіни 60 см, до бокових 120 см.

Відтворення бінауральної пари сигналів через телефони стимулює у слухача відчуття розташування джерела в позиції  $\Gamma_p$ . Проте це джерело звуку буде вже уявним. Змінюючи передавальні функції просторових фільтрів  $H_{\varphi^1}$  і  $H_{\varphi^2}$  у відповідності з положенням джерела звуку відносно голови слухача, можна за допомогою електричної схеми, також забезпечити локалізацію УЗД в будь-якому напрямку азимутальної площини. Тож сукупність просторових фільтрів  $\{H_{\varphi^1}\}$  і  $\{H_{\varphi^2}\}$  являють собою простий регулятор напрямків бінауральної звукової системи.

Попри значні успіхи, поки що не існує методів, які дають змогу усунути небажаний вплив реверберації приміщення, з одного боку, та розширити зону слухання – з іншого. Це, безумовно, стримує широке використання бінауральної системи звукопередачі під час запису, на телебаченні та радіомовленні. Одним з недоліків бінауральної стереофонії є відсутність повної сумісності з традиційною стереофонічною системою звукопередачі. Втім, бінауральна технологія має свою нішу і досить активно розвивається.

**Системи HRTF.** HRTF (Head Related Transfer Function) – це процес, за допомогою якого наші два вуха визначають місце розташування джерела. Наша голова та тулуб є, в деякій мірі, перешкодою, яка фільтрує і затримує звук. Тому вухо, яке є прихованим від джерела звуку, головою сприймає змінені звукові сигнали, які вже мозком інтерпретуються відповідним чином для визначення місця розташування джерела звуку. Звук, який сприймає наше вухо, створює на барабанну перетинку тиск. Для визначення тиску необхідно спочатку визначити характеристику імпульса сигналу від джерела звуку, тобто силу з якою звукова хвиля від джерела звуку впливає на барабанну перетинку. Ця залежність називається Head Related Impulse Response (HRIR), а її інтегральне перетворення Фур'є називається HRTF.

HRTF представляє собою надзвичайно складну функцію з чотирма змінними: три просторових координати і частота. За умов використання

сферичних координат для визначення відстані до джерела звуку більших, ніж один метр, вважається, що джерела звуку знаходяться у дальньому полі і значення HRTF зменшується обернено пропорційно відстані. Більшість вимірювань HRTF проводяться саме в дальньому полі, що істотно спрощує HRTF до функції азимута, висоти та частоти. Тобто відбувається спрощення завдяки позбавленню від четвертої змінної. Після цього використовують отримані значення вимірювань і в результаті, під час програвання звуку відтворюється з таким ж просторовим розташуванням, як і при природньому прослуховуванні. Техніка HRTF використовується вже декілька десятків років для забезпечення високої якості стереозаписів. Найкращі результати отримані при прослуховуванні запису в навушниках одним слухачем.

Навушники можуть мати провали та піки в своїх частотних характеристиках, які відповідають характеристикам вушної раковини. Якщо такої відповідності не існує, то і сприйняття звуку, джерело якого знаходиться у вертикальній площині, може бути погіршене. Тобто ми будемо слухати і чути лише звук від джерел, які знаходяться в горизонтальній площині.

За умов прослуховування в навушниках складається враження, що джерело звуку знаходиться дуже близько. І дійсно, фізичне джерело звуку розташовується дуже близько до вуха, тому необхідна компенсація для того, щоб позбавитись від акустичних сигналів, які впливають на визначення місця розташування фізичних джерел звуку.

Використання акустичних колонок (рис.1.7) дає змогу обійти більшість проблем, але при цьому не зовсім зрозуміло, яким чином можна використовувати колонки для відтворення бінаурального звуку (тобто звуку, який призначений для прослуховування в навушниках, коли частина сигналу призначена для одного вуха, а інша частина – для іншого). Як тільки ми підключаємо замість навушників колонки, наше праве вухо починає чути не тільки звук, призначений для нього, але і звуки, призначені для лівого вуха. Одним з рішень такої проблеми є використання техніки cross-talk-cancelled

stereo або transaural, яка частіше називається просто алгоритм stereo crosstalk cancellation (для зручності CC).

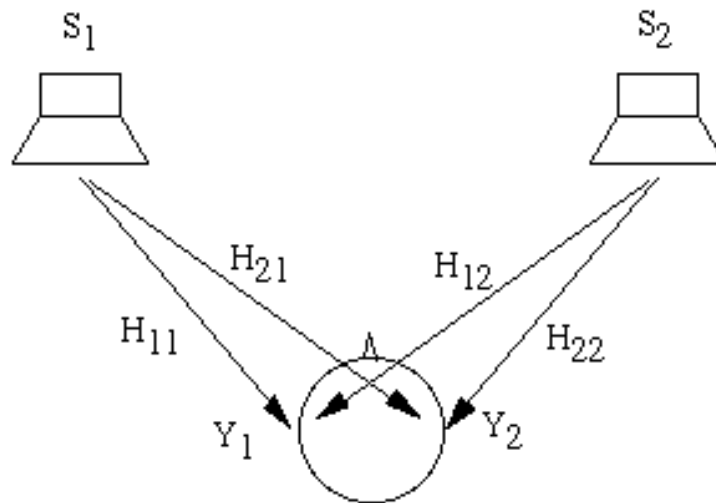


Рис 1.7. Схема відтворення колонками сигналів  $S_1$  та  $S_2$  і їх сприйняття людиною

Правильне сприйняття звучання досягається тільки в районі так званого «sweet spot», місце в якому і передбачалось місце розташування слухача. Тому те, як ми чуємо звук залежить не тільки від того, як був зроблений запис, але й від того, з якого місця між колонками ми слухаємо запис.

За умов грамотного використання алгоритмів CC можна отримати досить пристойні результати, які забезпечують сприйняття звуку, якщо джерела розташовані у вертикальній площині та горизонтальній. Фантомні джерела звуку можуть розташовуватися далеко поза межами лінійного сегмента між двома колонками.

Для створення переконливого 3D звучання достатньо двох звукових каналів. Головне - відтворити звуковий тиск на барабанні перетинки в лівому та правому вухах таким чином, ніби слухач знаходився в реальному звуковому середовищі.

Через те, що розрахунок HRTF функцій складна задача, в багатьох системах просторового звуку (spatial audio system) розробники покладаються на використання даних, отриманих експериментальним шляхом. Проте

основною причиною використання HRTF є бажання створити ефект elevation (звук в вертикальній площині), поряд з азимутальними звуковими ефектами. При цьому сприйняття звукових сигналів, джерела яких розташовані у вертикальній площині, є надзвичайно чутливим до особливостей кожного конкретного слухача. В результаті утворилось чотири методи розрахунку HRTF:

- **Використання стандартних функцій HRTF.** Такий метод забезпечує посередні результати відтворення ефектів elevation для деякого відсотку слухачів, але це найпоширеніший метод в недорогих системах. На сьогодні, ні IEEE, ні ACM, ні AES не визначили стандарт для HRTF.
- **Використання одного типу HRTF функцій з набору стандартних функцій.** В такому випадку необхідно визначити HRTF для невеликої кількості людей, які є представниками різних класів слухачів, і надати користувачу простий спосіб обрати саме той набір HRTF функцій, які найкращим чином відповідають йому (мається на увазі зріст, форма голови, розташування вух). Незважаючи на те, що такий метод запропонований, поки що ніяких стандартних наборів HRTF функцій не існує.
- **Використання індивідуальних HRTF функцій.** В такому випадку необхідно проводити визначення HRTF, виходячи з параметрів конкретного слухача, що сама по собі складна і затяжна процедура. Однак, ця процедура забезпечує найкращі результати.
- **Використання методів моделювання параметрів HRTF, які можуть бути адаптовані під кожного конкретного слухача.** Саме цей метод зараз застосовується скрізь в технологіях 3D звуку.

На практиці існують деякі проблеми, пов'язані зі створенням бази HRTF функцій за допомогою манекена. Результат буде відповідати очікуванням, якщо манекен і слухач мають голови однакових розмірів та форми, а також вушні раковини однакових розмірів та форми. Тільки за таких умов можна коректно відтворити ефект звучання у вертикальній

площині і гарантувати правильне визначення місце розташування джерел у просторі. Записи створені з використанням HRTF називаються *binaural recordings*, і вони забезпечують високоякісний 3D звук. Слухати такі записи необхідно в навушниках, до того ж бажано в спеціальних навушниках.

Такі записи коштують дорожче звичайних. Щоб коректно відтворити такі записи на колонках необхідно додатково використовувати техніку СС. Але головний недолік даного методу — відсутність інтерактивності. Без додаткових механізмів, які відстежують положення голови користувача, забезпечити інтерактивності неможливо. Існує навіть приказка, що використовувати HRTF для інтерактивного 3D звуку, це все одно, що використовувати ложку замість викрутки: інструмент не відповідає задачі.

## РОЗДІЛ 2. МЕТОДИ ОЦІНКИ РОЗБІРЛИВОСТІ МОВИ

### 2.1. Суб'єктивні методи оцінки розбірливості мови

Артикуляційні випробування проводяться згідно з ГОСТ 16600-72. В документі вказані класи якості розбірливості мови (табл. 2.1) і визначені види артикуляційних таблиць.

Таблиця 2.1. Класи якості розбірливості мови згідно ГОСТ 16600-72

Клас якості	Характеристика класу якості	Норми складової розбірливості мови для трактів з параметричним компадуванням	Норми складової розбірливості мови для трактів з кодуванням

			хвилі мовного сигналу
Вищий	Розуміння мови без найменшого напруження уваги	>93	>80
I	Розуміння мови без труднощів	86-93	56-80
II	Розуміння мови з напруженням без перепитувань і повторень	76-85	41-55
III	Розуміння мови з великим напруженням уваги з перепитуваннями і повторень	61-75	25-40
IV	Повна нерозбірливість зв'язного тексту (зрив зв'язку)	45-60	<25

Мінімально допустимим значенням  $D_0$  або  $W_0$  для кожного класу якості є найменшим значенням розбірливості.

В документі також вказані вимоги до приладів та обладнань, артикуляційних бригад, умов тренування членів артикуляційної бригади. Пояснюються умови проведення випробувань із застосуванням артикуляційних таблиць звукосполучень. Також пояснюються алгоритми обробки результатів експериментів.

За умов використання таблиць для артикуляційних випробувань необхідно дотримуватись певного алгоритму:



- Диктор зачитує чотири таблиці, використовуючи випробувальний тракт.
- Два-три слухачі записують. Для кожного аудитора визначають відсоток правильно прийнятих звукосполучень на кожну таблицю.
- Обраховується середнє значення розбірливості звуків  $D_{cp}$  та середньоквадратичне відхилення  $\sigma_D$ :

$$D_{cp} = \frac{1}{K} \sum_{i=1}^K D_i; \quad (1)$$

$$\sigma_D = \sqrt{\frac{1}{K-1} \sum_{i=1}^K (D_i - D_{cp})^2}; \quad (2)$$

де  $D_i$  - розбірливість одним слухачем однієї таблиці;  $K = m \cdot n$  - загальне число таблиць, сприйняті усіма слухачами;  $m$  - число слухачів;  $n$  - число взятих таблиць.

Якщо  $|D_i - D_{cp}| \geq 3\sigma_D$  то  $D_{cp}$  і  $\sigma_D$  обчислюють повторно, зменшуючи число вимірів.

Наступним кроком визначають нижню межу довірчого інтервалу з імовірністю 0.95:

$$D_n = D_{cp} - C_K \sigma_D; \quad (3)$$

де  $C_K$  - коефіцієнт, який враховує довірчу імовірність, визначається з таблиці 2.

Таблиця 2.2. Коефіцієнти для проведення розрахунків

К	С <sub>К</sub>	К	С <sub>К</sub>
2	4.47	14	0.47
4	1.17	16	0.44

6	0.82	18	0.41
8	0.67	20	0.39
10	0.58	24	0.35
12	0.52	28	0.30

Якщо  $D_n \leq D_0 \leq D_{cp}$  то вимірювання продовжують. В цьому випадку обчислюють відношення  $(D_{cp} - D_0) / \sigma_D$  і з таблиці 2 знаходять найближче значення  $C_K$  і відповідне значення  $K$  для цього відношення.

Існує також стандарт РФ ГОСТ 50840-95 «Передача речи по трактам связи» заснований на методах вимірювання показників розбірливості та якості мови, а також розпізнавання голосу диктора. Стандарт регламентує одержання комплексної оцінки якості передачі мовлення. На траси телефонного провідного і радіозв'язку, в яких використовується аналоговий мовний сигнал, поширюється даний стандарт. Також на прилади, що містять перетворювачі мовного сигналу у цифровий варіант.

Комплексна оцінка якості передачі мови включає в себе:

- оцінку складової розбірливості;
- показник якості мови, отриманий методом оцінки шести селективних ознак;
- показник якості мови, отриманий методом парних порівнянь з контрольним трактом;
- метод парних порівнянь для оцінки розбірливості і розпізнавання голосу диктора;
- вимірювання фразової розбірливості при прискореному темпі вимови.

В складі даного стандарту додається 1000 складових артикуляційних таблиць і 38 таблиць тестових фраз.

Відносна кількість вірно зрозумілих елементів мови — розбіливість мови — оцінюється якість мовлення, яка є суб'єктивною оцінкою звучання мови у випробувальному тракті:

- у порівнянні зі звучанням у контрольному тракті;
- порівняно зі звучанням в іншому тракті.

Класи якості і норми розбірливості мови дещо відрізняються від прийнятих в ГОСТ 16600-72 (табл. 2.3).

Таблиця 2.3. Класи якості розбірливості мови згідно ГОСТ 50840-95

Клас якості	Характеристика класу якості	Норми складової розбірливості мови для трактів з параметричним компадуванням	Норми складової розбірливості мови для трактів з кодуванням хвилі мовного сигналу
Вищий	Розуміння мови без найменшого напруження уваги	>93	>80
I	Розуміння мови без труднощів	86-93	56-80
II	Розуміння мови з напруженням без перепитувань і повторень	76-85	41-55
III	Розуміння мови з великим напруженням уваги з перепитуваннями і повторень	61-75	25-40

IV	Повна нерозбірливість зв'язного тексту (зрив зв'язку)	45-60	<25
----	---	-------	-----

Склад артикуляційної бригади: не менше трьох попередньо тренованих дикторів і трьох аудиторів віком 18-30 років.

Зразок артикуляційних таблиць наведено в таблиці 2.4

Таблиця 2.4. Артикуляційні таблиці

з <sup>в</sup> уф <sup>ь</sup>	ст <sup>ь</sup> ч <sup>ь</sup>	ж <sup>ь</sup> ен	ер	н <sup>ь</sup> я <sup>ь</sup> р <sup>ь</sup>	зн <sup>ь</sup> ун	га <sup>ь</sup> н <sup>ь</sup>	ко <sup>ь</sup> с <sup>ь</sup>	сы <sup>ь</sup> м	ло <sup>ь</sup> ц
б <sup>ь</sup> ю <sup>ь</sup> н	хо <sup>ь</sup> л <sup>ь</sup>	дя <sup>ь</sup> ф <sup>ь</sup>	не <sup>ь</sup> н <sup>ь</sup>	ш <sup>ь</sup> то <sup>ь</sup> т <sup>ь</sup>	лю <sup>ь</sup> п <sup>ь</sup>	про <sup>ь</sup> ц	ши <sup>ь</sup> ч	пя <sup>ь</sup> щ	ще <sup>ь</sup> т <sup>ь</sup>
ц <sup>ь</sup> у <sup>ь</sup> ч	ки <sup>ь</sup> к	мо <sup>ь</sup> т <sup>ь</sup>	дре <sup>ь</sup> т <sup>ь</sup>	рю <sup>ь</sup> х	че <sup>ь</sup> с <sup>ь</sup>	ми <sup>ь</sup> ф <sup>ь</sup>	тра <sup>ь</sup> р	зя <sup>ь</sup> к	но <sup>ь</sup> т <sup>ь</sup>
пл <sup>ь</sup> ы <sup>ь</sup> к	фу <sup>ь</sup> ч	ч <sup>ь</sup> т <sup>ь</sup> э <sup>ь</sup> т	ва <sup>ь</sup> к	ле <sup>ь</sup> н	ту <sup>ь</sup> х	ву <sup>ь</sup> р	зо <sup>ь</sup> с	хе <sup>ь</sup> т	зда <sup>ь</sup> л
се <sup>ь</sup> ф	ка <sup>ь</sup> т <sup>ь</sup>	до <sup>ь</sup> р	по <sup>ь</sup> н	тя <sup>ь</sup> ч	су <sup>ь</sup> с <sup>ь</sup>	ба <sup>ь</sup> н	ве <sup>ь</sup> н <sup>ь</sup>	ти <sup>ь</sup> с <sup>ь</sup> т <sup>ь</sup>	ру <sup>ь</sup> л

ре <sup>ь</sup> к	ни <sup>ь</sup> щ	зн <sup>ь</sup> ом	ры <sup>ь</sup> с	ге <sup>ь</sup> к	та <sup>ь</sup> н <sup>ь</sup>	ву <sup>ь</sup> ф	тре <sup>ь</sup> п	ме <sup>ь</sup> х	па <sup>ь</sup> с <sup>ь</sup>
зы <sup>ь</sup> р <sup>ь</sup>	ч <sup>ь</sup> т <sup>ь</sup> уй	бы <sup>ь</sup> с	тет	ске <sup>ь</sup> с	мо <sup>ь</sup> ф <sup>ь</sup>	ля <sup>ь</sup> м	го <sup>ь</sup> ф	ся <sup>ь</sup> х	ро <sup>ь</sup> ц
фо <sup>ь</sup> к	шу <sup>ь</sup> с <sup>ь</sup>	ху <sup>ь</sup> л	шта <sup>ь</sup> й	ми <sup>ь</sup> с	два <sup>ь</sup> т	три <sup>ь</sup> ю <sup>ь</sup> ль	ну <sup>ь</sup> ц	ре <sup>ь</sup> ль	ши <sup>ь</sup> й
да <sup>ь</sup> м <sup>ь</sup>	пю <sup>ь</sup> м	ка <sup>ь</sup> ц	бе <sup>ь</sup> к	чи <sup>ь</sup> н <sup>ь</sup>	щу <sup>ь</sup> рь	лу <sup>ь</sup> ф	фи <sup>ь</sup> ш	су <sup>ь</sup> зн <sup>ь</sup>	пре <sup>ь</sup> н <sup>ь</sup>
ста <sup>ь</sup> р	ба <sup>ь</sup> н <sup>ь</sup>	вю <sup>ь</sup> р	на <sup>ь</sup> х	бе <sup>ь</sup> м	са <sup>ь</sup> х	де <sup>ь</sup> с <sup>ь</sup>	жу <sup>ь</sup> с	цо <sup>ь</sup> й	ло <sup>ь</sup> рь

Для вимірювання артикуляційної розбірливості використовується 5 складових таблиць.

Обробка результатів передбачає обчислення середнього значення і середньоквадратичного відхилення оцінки розбірливості:

$$S = \frac{1}{N} \sum_{i=1}^N S_i \quad (4)$$

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (S_i - S)^2} \quad (5)$$

де  $S_i$  - результат одиничного вимірювання;  $N$  - кількість вимірювань.

Потім відсіюються сумнівні дані, а також відкидаються результати, для яких  $S_i - S \geq 2\sigma$ . Після цього обчислюється нове середнє значення:

$$S = \frac{1}{N - k} \sum_{i=1}^{N-k} S_i \quad (6)$$

де  $k$  – число відкинутих вимірювань.

## 2.2. Об'єктивні методи вимірювання розбірливості мови

На сьогодні існує значна кількість методів вимірювання та розрахунку розбірливості мови, які застосовуються для оцінки акустики приміщення ті ліній зв'язку. Методи поділяються на декілька груп:

- формантні;
- теоретико-інформаційні;
- модуляційні;
- емпіричні.

Теорією розбірливості мови займалися свого часу - Ю.С.Биков, М.А.Сапожков, Н.Б.Покровський. Автори надавали перевагу формантному методу.

Іноземні автори працювали над наступною теорією, вже третьою. Вона забезпечувала гарний збіг розрахунку та експерименту. Проте Ю.Биков і Н.Покровський не рекомендують застосовувати дану версію на практиці оскільки вона є досить складною і трудомісткою, а також не зовсім підтверджується дослідями.

К.Крітер представив наступну теорію – індекс артикуляції. Теорія також була четвертою версією формантного методу.

На сьогодні існує п'ять версій модуляційного методу: STI, STIr, RASTI, STITEL, STIPA.

Модуляційний метод був створений в 70х роках минулого століття. Т.Ходкаст і Г.Стінкен запропонували використовувати тестовий сигнал у вигляді шуму. Шум модульований за амплітудою періодичним сигналом. Ця ідея виявилась досить вдалою, оскільки дала змогу враховувати вплив шумової та ревербераційної завади на розбірливість мови.

Згодом Ю.К.Калинцев запропонував теоретико-інформаційний метод. Розрахований цей метод на експертизу вокодерних ліній зв'язку. Таким чином даний метод дає змогу проводити експертизу цифрових ліній зв'язку, на відміну від формантного і модуляційного методів.

Емпіричні методи на сьогодні є досить актуальними, незважаючи на ряд суттєвих недоліків у вигляді низької точності отримуваних недоліків. Проте їх значною перевагою є простота і швидкість обчислень.

### **РОЗДІЛ 3. ЛОКАЛІЗАЦІЯ УЯВНОГО ДЖЕРЕЛА ЗВУКУ В ОКТАВНИХ СМУГАХ ЧАСТОТ СЛУХОВОГО ДІАПАЗОНУ**

#### **3.1. Локалізація уявного джерела звуку за взаємною кореляційною функцією стереопари**

Кореляційні моделі механізму локалізації. В загальному випадку місцезрешування УДЗ в суб'єктивному слуховому просторі не співпадає з розташуванням дійсного джерела звуку. З цієї точки зору певною перевагою володіють кореляційні та асоціативні моделі просторового слуху.

Кореляційні моделі мають досить тісний зв'язок з нейрофізіологічними даними. Вони пояснюють механізм оцінки бінауральних параметрів  $\Delta L_\epsilon$  і  $\Delta \tau_\epsilon$ , а також феномен створення УДЗ при стереовідтворенні.

Розташування максимуму взаємної кореляційної функції пов'язане з боковим зміщенням джерела звуку. Максимальне значення функції та її ширина – зі ступенем когерентності бінауральних сигналів. Утворення УДЗ стає можливим, якщо величини коефіцієнту кореляції сигналів, які сприймаються від різних джерел, перевищують порогове значення.

Проте все ще неможливо розрахувати положення УДЗ на лінії бази гучномовців. Цей недолік можна подолати, зробивши наступні припущення:

- оцінка часового зсуву сигналів визначається за величиною  $\Delta\tau$ , при якій функція взаємної кореляції сигналів сягає свого максимального значення; в результаті цього процесу часового «сканування» часовий зсув вхідних сигналів компенсується в слуховій системі, коли функція взаємної кореляції сягає максимального значення;
- часовий зсув перетворюється в еквівалентну різницю рівнів шляхом ослаблення по інтенсивності сигналів, що спізнюються; це відбувається внаслідок бокового зсуву;
- напрямок на джерело звуку співпадає з положенням максимуму функції взаємної кореляції бінауральних сигналів в суб'єктивному слуховому просторі;
- максимальне значення функції взаємної кореляції одночасно пов'язане з рівнем та зі значенням  $\Delta\tau_0$  бінауральних сигналів.

**Асоціативна модель сприйняття звукових образів.** Асоціативна модель локалізації передбачає наявність двох послідовних етапів обробки інформації в слуховій системі: асоціація місця дійсного джерела звуку в просторі і асоціація форми, де можливо утворення УДЗ і оцінка азимута.

Проблема локалізації звукових образів у просторі трактується наступним чином: одинарне дійсне джерело звуку завжди викликає одну асоціацію місця, яка визначає його розташування у просторі; два рознесені джерела звуку на етапі асоціації місця розділені одне від одного. Якщо

сигнали джерел некорельовані, то на етапі розпізнавання форми вони викликають дві асоціативні форми, сприймаються як два окремих звукових образи, розташування у просторі яких, як і раніше визначаються за принципом локалізації місця на першому етапі обробки інформації. Наявність кореляції між сигналами джерел на етапі асоціації форми призводить до злиття процесів слухання і утворення УДЗ. В цьому випадку випадку локалізація УДЗ є наслідком двох етапів обробки інформації – асоціації місця і форми.

### 3.2. Порівняння ефективності локалізації імпульсного і мовного сигналів в октавних смугах частот

**Кореляція сигналів в частотних смугах.** Кореляція імпульсного сигналу закономірно знижується зі зростанням середньої частоти смуги; в мові – кореляція сигналу в частотних смугах несе нерівномірний характер. На частоті 1000 Гц спостерігається зміна класичного вигляду функції кореляції. Загальна кореляція сигналу знижується.

У центральних точках, ступінь кореляції вище та по смугам більш рівномірний. Проте, це не може слугувати єдиною ознакою для вимірювання коефіцієнта крос-кореляції. Необхідно враховувати зміщення максимуму та її загальний вигляд.

На високих частотах графіки стають менш симетричними відносно ординати, це вказує на підвищення просторовості та об'ємності звуку, що в свою чергу негативно впливає на розбірливість. Також на високих частотах різниця між функціями кореляції імпульсного сигналу та мовного стає більш помітною. На високих частотах кореляція має багато максимумів і стає неоднозначною. Більш інформативною є огинаюча, а не сама функція.

Підтверджується загальна закономірність хорошої кореляції низькочастотного сигналу і погіршення кореляції з підвищенням частоти.

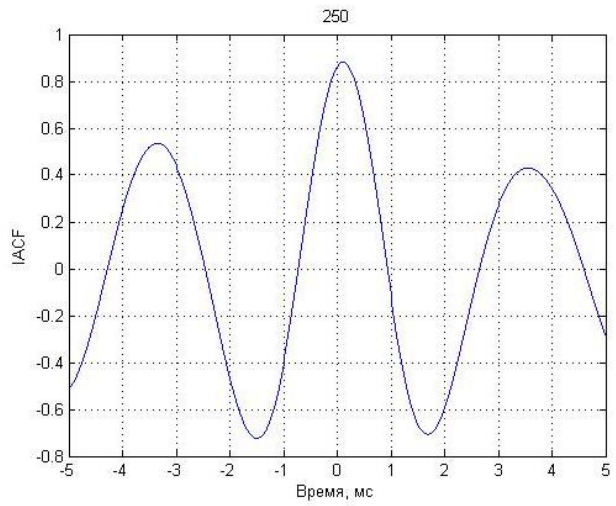
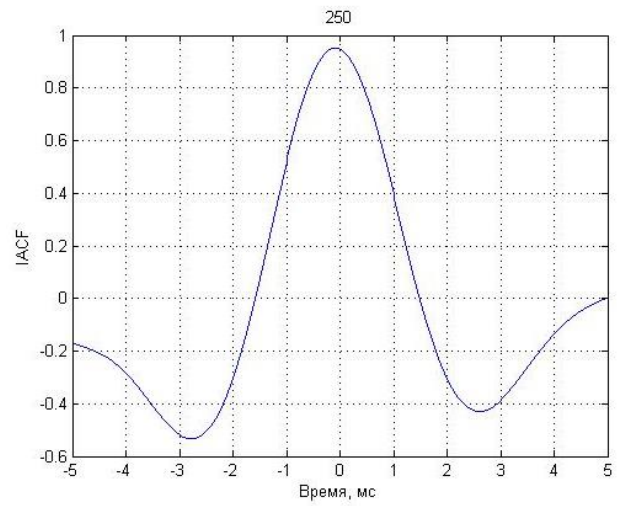
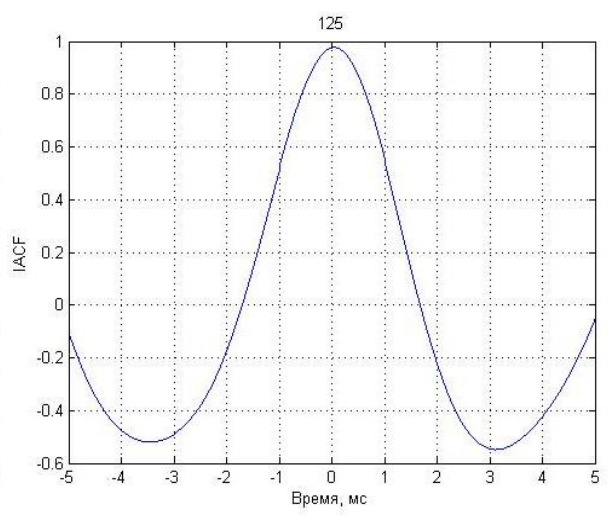
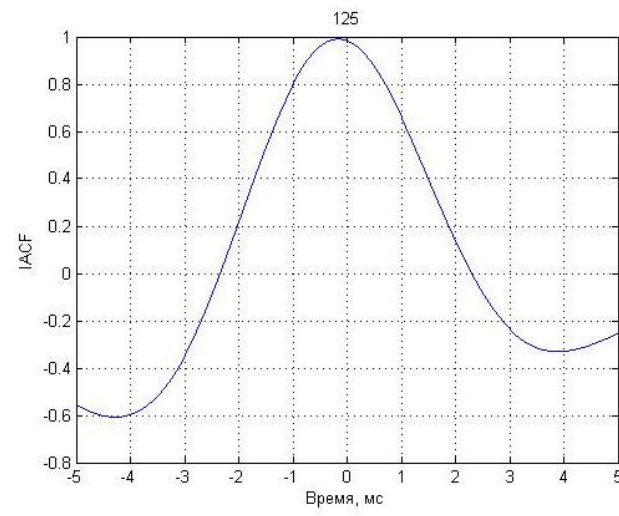
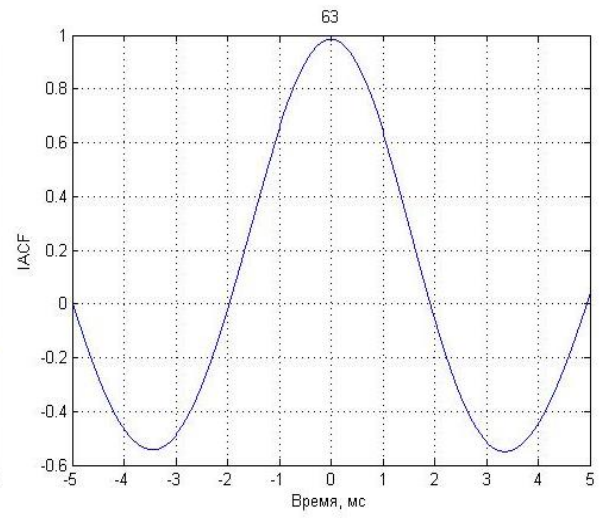
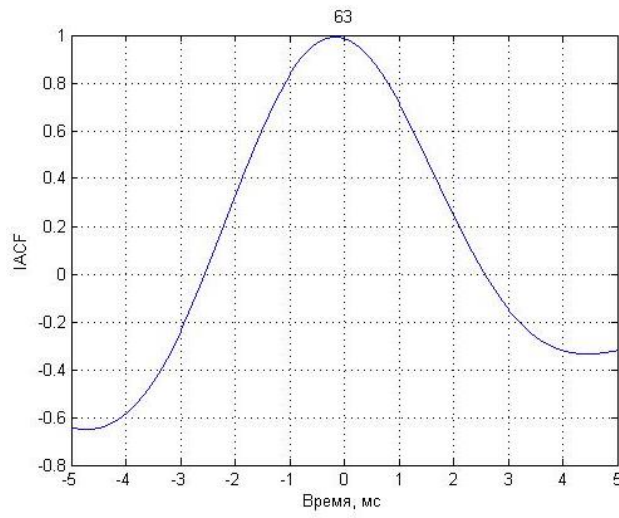


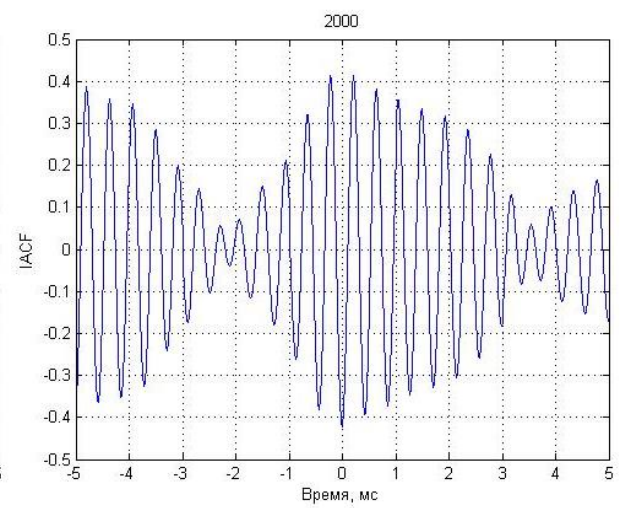
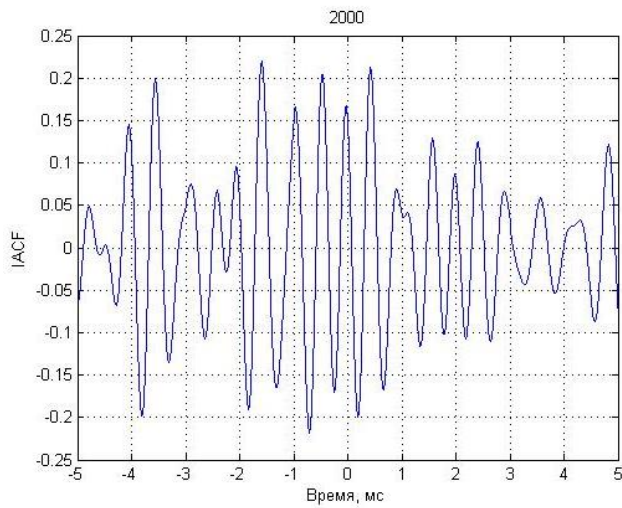
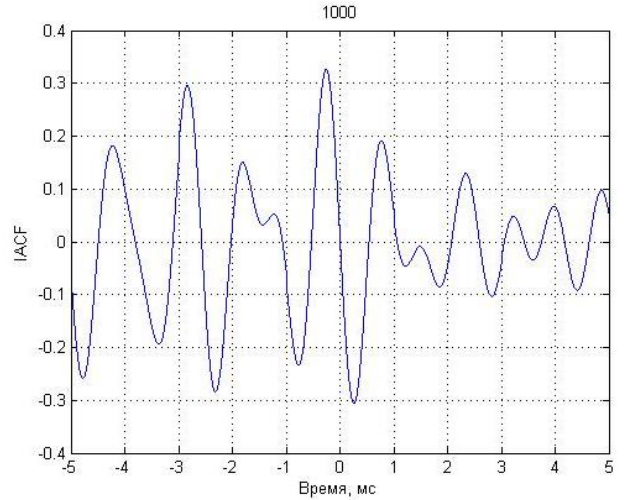
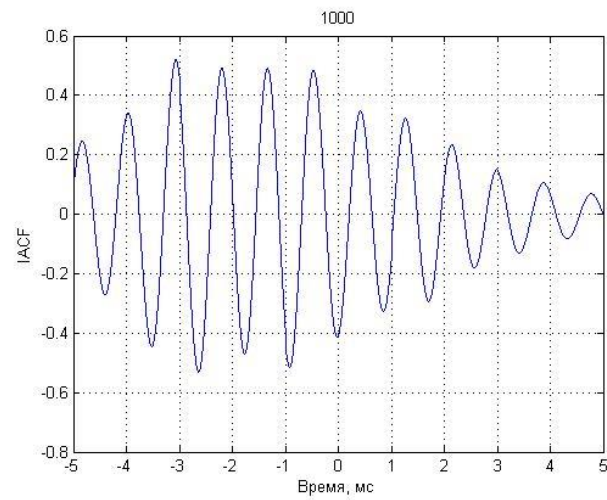
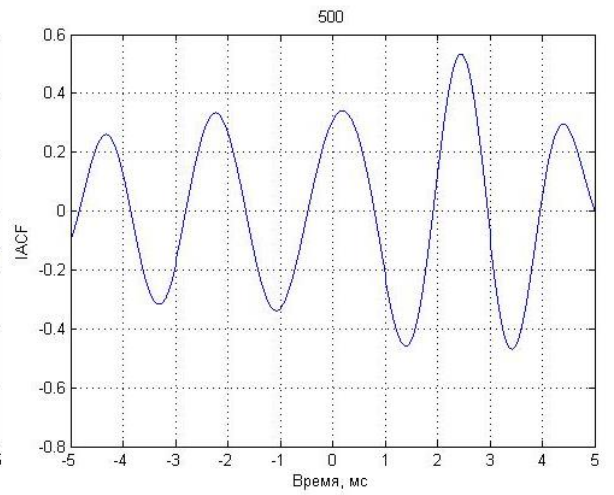
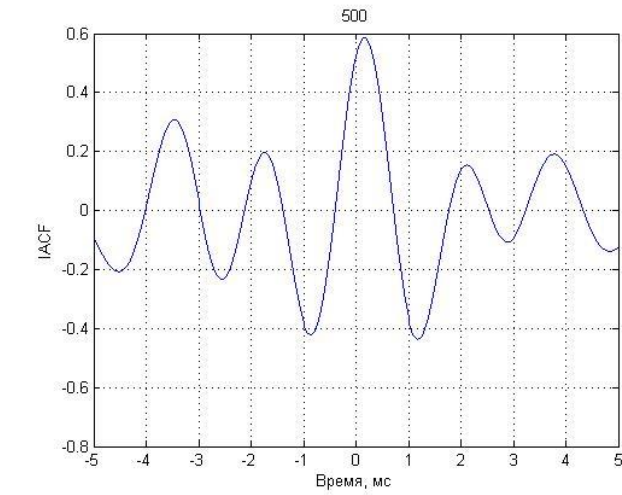
У різних смугах частот положення УЗД може бути різним і це ускладнює локалізацію і як наслідок розбірливість.

Недостатньо аналізувати функцію крос-кореляції на інтервалі  $[-1;1]$ , варто звернути увагу на загальний вид функції. Можна помилково охарактеризувати приміщення, пропустивши головний максимум при поганій кореляції і прийняти отриманий графік, як хороший результат з відмінною кореляцією. Для мовних сигналів кореляція має досить нерівномірний характер. Прослідковується різке зростання кореляції на частоті 1000 Гц. Загальна кореляція сигналу знижується. На високих частотах графіки стають менш симетричними відносно ординати, це вказує на підвищення просторовості та об'ємності звуку, що в свою чергу негативно впливає на розбірливість. Більш помітні спотворення спектру мовних сигналів, оскільки спектр імпульсного сигналу має простішу структуру, на відміну від мовного сигналу. Графіки спектрів у смугах стандартних частот для мови додаток А.

Імпульс

Мова





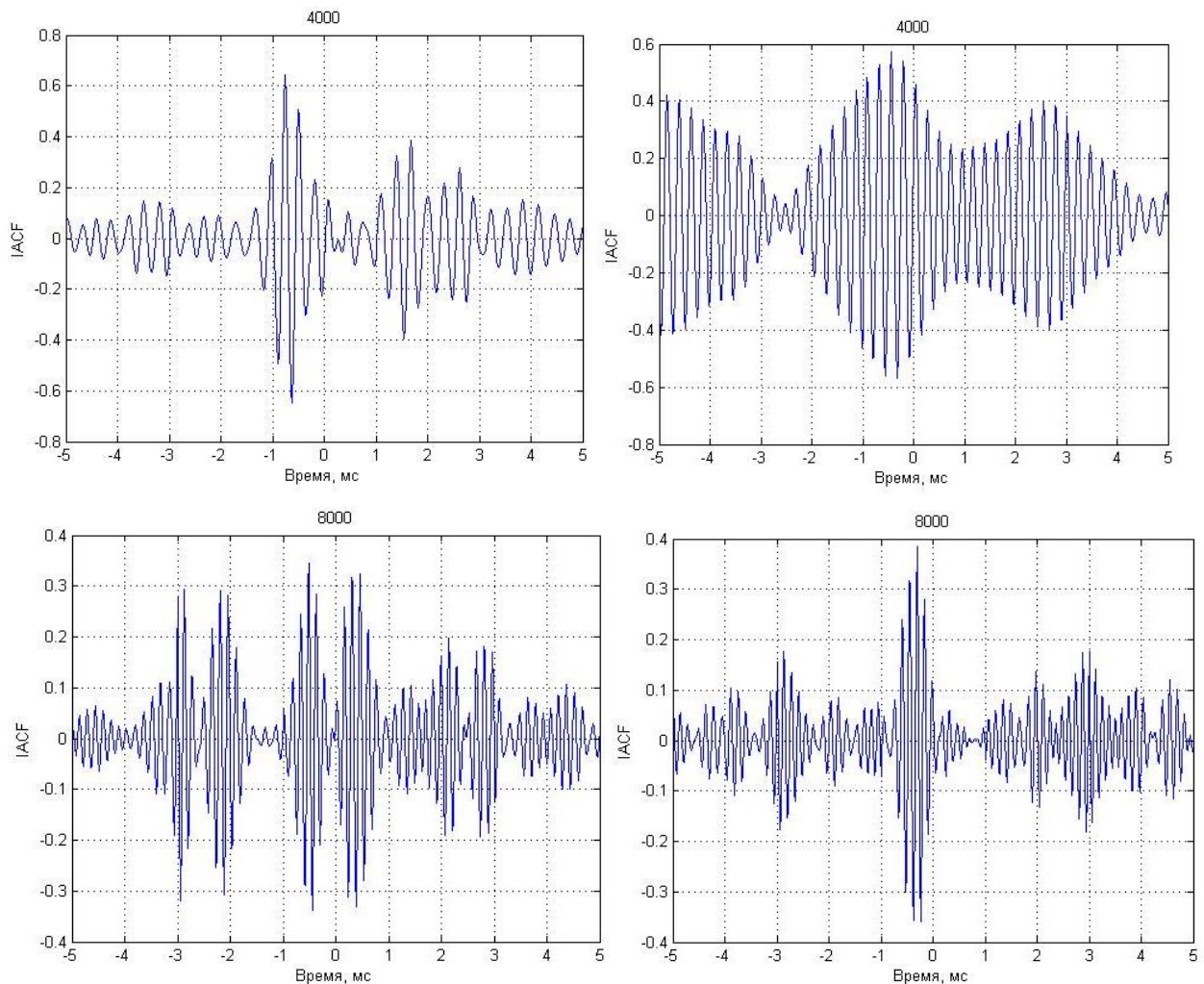


Рис.4.1 Крос-кореляційна функція в октавних смугах частот

#### РОЗДІЛ 4. ОЦІНКА РОЗБІРЛИВОСТІ МОВИ ЗА КОЕФІЦІЄНТОМ МІЖВУШНОЇ КОРЕЛЯЦІЇ МОВНОГО СИГНАЛУ

Нині існує безліч об'єктивних методів оцінки розбірливості мови в акустичних трактах та приміщеннях. Широкопasmовий шум, у ролі досліджуваного сигналу, використовують майже у всіх цих методах. Широкопasmовий шум має чималий набір слухових частот, але, як показує практика, досить сильно відрізняється від мови в плані сприйняття їх людським організмом. Мовний сигнал має складнішу структуру, порівняно з шумом. Акустичні коливання повітряного струменя є фізичною основою

мовного звучання, які формуються органами артикуляції. Втім не лише акустика звучання визначає те, як людина чує мову. Фонетична система мови відіграє вирішальну роль, впливаючи на механізми і результат сприйняття мови людським організмом на слух. Саме цим пояснюються помилки при сприйнятті усної мови в усіляких ситуаціях спілкування.

Складова і фразова розбірливості мови, якість мови в порівнянні з якістю мови еталонного тракту та якість мови в реальних умовах роботи є основними параметрами якості мови. Досить складно дослідити, оцінити та проаналізувати такі параметри за допомогою широкосмугового шуму, оскільки в його складі не маємо ні фраз, ні складів. Тому доречніше та достовірніше використовувати мовний сигнал для об'єктивної оцінки розбірливості мови.

Один з методів оцінки розбірливості мови в приміщенні — використання коефіцієнтів міжвушної кореляції, одержаних шляхом вимірювань бінауральних мовленнєвих сигналів. Тобто додатково вимірюється ще один об'єктивний параметр, який характеризує сприйняття звуку аудитором, а саме функція міжвушної кореляції (IAFC) [6]. За максимальним значенням функції, так званим коефіцієнтом міжвушної кореляції (IACC), з'являється спроможність оцінити розбірливість мови в приміщенні

За значеннями звукового тиску на обох вухах визначається коефіцієнт кореляції. Він демонструє ступінь подібності сигналів. Наприклад, якщо сигнали на двох вухах однакові, то джерело звуку буде здаватися точковим, якщо ж сигнали, які надходять до лівого та правого вух - різні, це створює враження розширення звукового джерела, тобто коефіцієнт кореляції в певних межах зменшується.

Кореляційна функція нормована на взаємну енергію відповідних сигналів:



$$IACF(\tau) = \frac{\int_{-\infty}^{\infty} p_l(t) p_r(t+\tau) dt}{\sqrt{\int_{-\infty}^{\infty} p_l^2(t) dt \cdot \int_{-\infty}^{\infty} p_r^2(t) dt}}, \quad (7)$$

де  $\tau$  – часова затримка,  $p_l(t)$  та  $p_r(t)$  – звуковий тиск на лівому та правому вухах.

Визначається сам коефіцієнт міжвушної кореляції як:

$$IACC = \max |IACF(\tau)|, \text{ при } -1\text{мс} \leq \tau \leq 1\text{мс} \quad (8)$$

тобто за максимальним значенням функції, при зміні від -1мс до +1 мс.

Індекс чіткості для мови за імпульсним сигналом, який визначається наступним чином:

$$C_{50} = 10 \lg \left[ \int_0^{50\text{мс}} p^2(t) dt / \int_{50\text{мс}}^{\infty} p^2(t) dt \right]$$

Де  $p^2(t)$  енергія імпульсного сигналу

Обробка результатів вимірювань дозволяє знайти параметр, що називається «чіткістю»:

$$D := \int_0^{50\text{мс}} p^2(t) dt / \int_{50\text{мс}}^{\infty} p^2(t) dt$$

Через який можливо знайти коефіцієнт складової розбірливості

$$k_c = \sqrt[5]{D} = \sqrt[5]{1 / (1 + 10^{-C_{50}/10\text{дБ}})}$$

Експеримент проводиться в аудиторії 412 для моно- і стереосистеми озвучення залу. Геометрія приміщення і розташування точок вимірювання і стереосистеми представлені на рисунку 4.1. Вимірювання проводилися за допомогою штучної голови. Розміри залу 14,3х18,5х6,25 м, об'ємом  $V \approx 1653\text{м}^3$ . Випромінювався імпульсний сигнал тривалістю 5 мс.

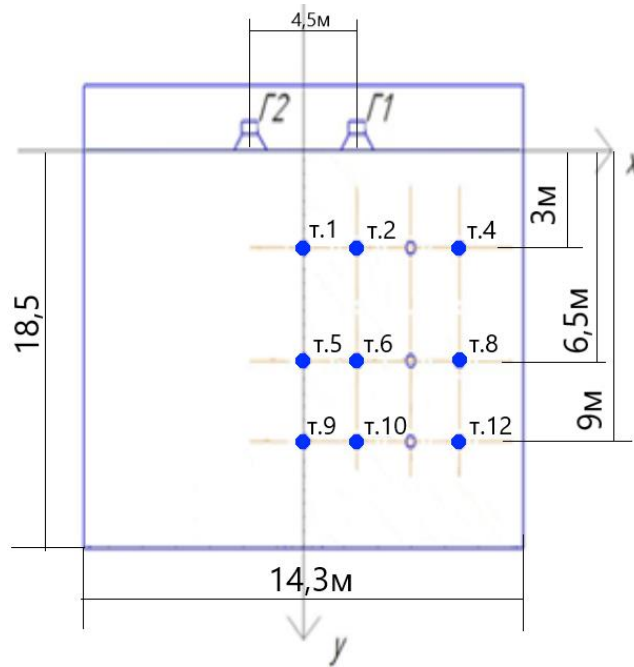


Рис 4.1 Геометрія розташування стереосистеми та точок вимірювання в залі

Експеримент проводився згідно ISO 3382-1:2009. Використовувались гучномовці Radiotekhnika S-30B, підсилювач потужності Park Audio V4-900 MkII, звукова карта M-Audio Fast Track Pro, манекен голови людини з вмонтованими вимірювальними мікрофонами Behringer ECM8000. Досліджувані точки обрані відповідно до місць розташування аудиторів.

Значення  $C_{50}$  більше нуля. Це свідчить про добру розбірливість мови. Як і слід було очікувати, значення  $C_{50}$  для стереосигналу вищі, ніж для мовного (більший вклад прямого звуку), особливо в перших рядах. Завеликі значення  $C_{50}$  в першому ряді свідчать про підвищену сухість мови. Набільші значення індексу чіткості в кожному ряду відповідають місцю навпроти гучномовця.

За представленими даними  $C_{50}$  наступним кроком розраховані відповідні значення коефіцієнту складової розбірливості. Значення  $K_s$  показані на шкалі класів розбірливості згідно ГОСТ Р 50840-95. Відповідна таблиця розподілу складової артикуляції представлена на таблиці 4.1. Оцінки розбірливості у вигляді відмінно, добре і т.д. присвоєні авторами.





об'єктивізованим (об.) способами, у відсотках

Для цього запропонована шкала відповідності цих двох величин, виходячи з класів розбірливості мови. За цією шкалою  $IACC > 0,5$  – розбірливість добра;  $0,3 < IACC < 0,5$  – задовільна;  $0,2 < IACC < 0,3$  – слабка;  $IACC < 0,2$  – незадовільна.

Ця величина значно більш корельована з відсотком артикуляції. Значення знаходяться в однакових класах розбірливості і відповідають оцінкам задовільно і погано.

Пунктиром показана суб'єктивна оцінка розбірливості виставлена слухачами за п'ятибальною системою (5 б.-100%).

#### Висновки:

Для оцінки розбірливості мови в залі можливо використовувати значення коефіцієнтів міжвушної кореляції мовного сигналу, одержаних шляхом вимірювання бінауральних сигналів на слухацьких місцях.

Вимірювання розбірливості мови в залі можливо проводити об'єктивним методом без артикуляційних вимірювань, тобто без участі аудиторів, що значно спрощує процедуру.

#### ВИСНОВКИ

В ході виконання магістерської роботи було встановлено, що крос-кореляційну функцію доцільно використовувати для аналізу бінаурального сприйняття. Спираючись на її вигляд можна доволі точно описати якість звучання в приміщенні через суб'єктивні ознаки, такі як ширина звукової панорами, об'ємність звучання, прозорість і чіткість звучання, багатство тембру тощо.

За допомогою функції крос-кореляції можна визначити напрямок до положення джерела звуку чи уявного джерела звуку, якщо у залі використовується стереовідтворення, при умові, що слухач знаходиться у зоні стереофонії. Якщо ж людина знаходиться за межами, вважають, що

настає «розрив», УЗД зникає, людина чує кожний гучномовець окремо. Проте у ході роботи було виявлено, що ближній до слухача гучномовець приймає властивості УЗД, а звук іншого гучномовця сприймається на ряду з відбитим звуком ближнього, утворюється так зване «злиття». Це викликано розмірами залу, кількістю гучномовців, відстанню між ними, загалом, умовами експерименту.

Якщо аналізувати крос-кореляцію у смугах частот (в даному випадку, октавних), то по вигляду функції можна оцінити розбірливість мови у приміщенні. Частотна складова в двох сигналах не однакова по значущості для розбірливості мови. Наприклад, для мови кореляція сигналу в частотних смугах несе нерівномірний характер (різке зростання кореляції на частоті 1000Гц). Загальна кореляція сигналу знижується. На високих частотах графіки стають менш симетричними відносно ординати, це вказує на підвищення просторовості та об'ємності звуку, що в свою чергу негативно впливає на розбірливість.

Також при дослідженні різних типів сигналів було виявлено, що показники якості звучання погіршуються при віддаленні від центру приміщення що цілком пояснюється властивостями та виглядом зони стереофонії. Проте у крайніх точках показники трохи підвищуються, що пояснюється близьким розташування досліджуваного манекену до бортів.

Слід зауважити, що функція по різному реагує на зміну сигналу, тобто для одного і того ж приміщення в одній і тій самій точці для різних типів сигналів функція має неоднаковий вигляд. Наприклад, в представлених результатах, це означає, що для симфонічної музики ця аудиторія має гарні показники, проте для мови необхідно вжити деяких заходів, щоб покращити розбірливість та загальне сприйняття інформації. Цікаво, що при відтворенні хорового співу показники є незадовільними. Це можна пояснити наявністю мовної складової у творі.

В цілому, після проведеної роботи можна зробити висновок, що використовуючи крос-кореляційну функцію можна оцінювати якість звучання у приміщеннях та робити прогнози щодо ймовірного покращення акустичних властивостей.

Перевагами даного методу є відносна простота та відповідність записаних результатів за допомогою манекену голови людини до сприйняття звуку реальним слухачем.

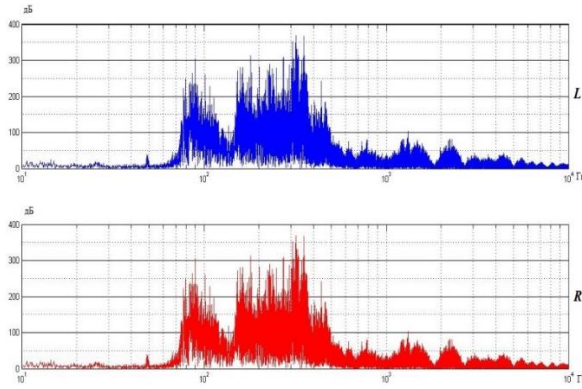
## СПИСОК ЛІТЕРАТУРИ

1. The Technology of binaural listening, Jens Blauert Editor
2. Ковригин С.Д. Архитектурно-строительная акустика: Учеб. Пособие для вузов. – М.: Высш. Школа, 1980. – 184с., ил.
3. THE EVOLUTION OF NEWER AUDITORIUM ACOUSTICS MEASURES J.S. Bradley Institute for Research in Construction National Research Council Ottawa, Canada, KIA OR6
4. Красильников В.А., Крылов В.В. Введение в физическую акустику. М., 1984.
5. Ухо и механизм восприятия звука. [Электронный ресурс] ООО «КМ онлайн» 1999-2014 <http://www.km.ru/zdorove/encyclopedia/ukho-i-mekhanizm-vospriyatiya-zvuka>
6. Цвинер Психоакустика – 2004
7. Ананьев Б. Г. Теория ощущений. — Л., 1961. — С. 579. — 928 с.
8. Бинауральная запись. [Электронный ресурс] Википедия 2014 [http://ru.wikipedia.org/wiki/Бинауральная\\_запись](http://ru.wikipedia.org/wiki/Бинауральная_запись)
9. Стенлі А. Гельфанд Слух: Введение в психологическую и физиологическую акустику/ О. К. Федоровой, О. П. Токарева - М.: Медицина - 1984 - 352с.
10. Бинауральная запись [Электронный ресурс] Журнал «Дом солнца» 2009 <http://www.sunhome.ru/journal/520775>
11. Бинауральные записи [Электронный ресурс] Звуки улиц Сергиева Посада 2012 [<http://www.oontz.ru/>]
12. Бинауральная стереофония. Современное состояние исследований [Электронный ресурс] Шоу-мастер 2004 [http://www.show-master.ru/categories/binauralnaya\\_stereofoniya\\_sovremennoe\\_sostoyanie\\_issledovaniy.html](http://www.show-master.ru/categories/binauralnaya_stereofoniya_sovremennoe_sostoyanie_issledovaniy.html)
13. Highest fidelity binaural recording in the world [Электронный ресурс] Binaural Airwaves 2013 <http://www.binauralairwaves.com/>
14. ISO 3382-1 Acoustics -- Measurement of room acoustic parameters -- Part 1: Performance spaces, 2010.

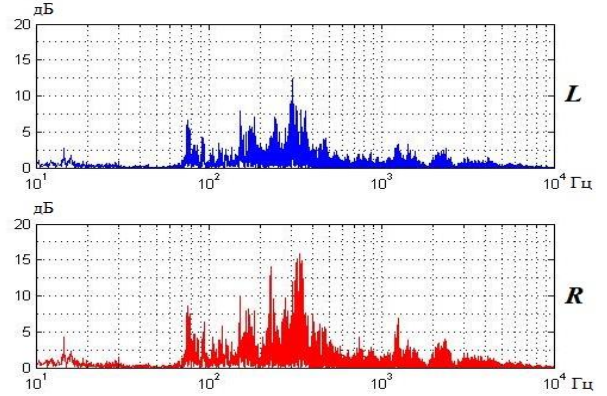
- 15.Лозенко А.А. Исследование пространственного звукового восприятия на основе бинауральных характеристик слуха / А.А. Лозенко, С.А. Луньова // Електроніка 2013. Міжнародна конференція молодих вчених: збірник статей. - К. 2013
- 16.Ю.А. Ковалгин Стерефония – М.: Радио и связь – 1989
- 17.Фурдуев В. В. Ф 95 Стерефония и многоканальные звуковые системы. М., «Энергия», 1973. 112с.
- 18.Электроакустика и звуковое вещание / под ред. Ю.А. Ковалгин. – М.: Горячая линия – Телеком, Радио и связь, 2007. – 872с.
- 19.Основы техники звукоусиления / Анерт В., Райхардт В. – М.: Радио и связь, 1984. – 319 с.
- 20.Алдошина И.А., Приттс Р. Музыкальная акустика. – С.-П.: Композитор, 2006.-719 с.
- 21.Вахитов Я. Ш. Теоретические основы элетроакустики и электроакустическая аппаратура. - М.: Искусство, 1982. -415с
- 22.Измерительный микрофон Behringer ECM8000 [Электронный ресурс] MusicLife UA <http://www.musiclife.kiev.ua/shop/studiyynyiy-mikrofon-behringer-ecm8000.html>
- 23.Расчет количества и выбор мощности включения звуковых и речевых оповещателей. журнал "Алгоритм безопасности" N6 2003г.
- 24.Акустика студий.,Михаил Ланэ., Архив журнала «Звукорежиссер»
- 25.Луньова С.А., Дідковський В.С. Основи архітектурної та фізіологічної акустики.-К.: Постприм, 2001.-420 с.
- 26.Beranek L. Music, Acoustic and Architecture. – N.Y.: Wiley&Sons, 1996.
- 27.Дідковський В.С., Луньова С.А., Богданов О.В. Архітектурна акустика.-К.

# Спектри мови у різних точках

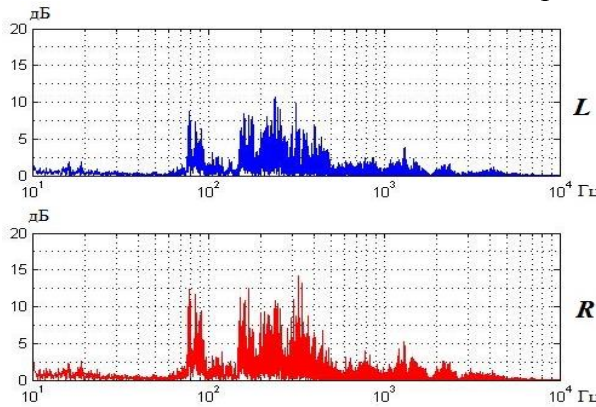
Вхідний спектр



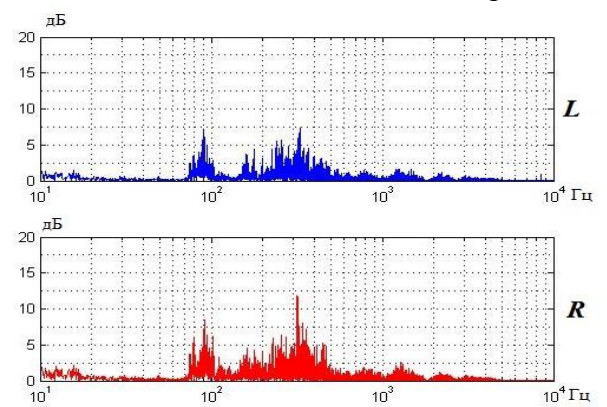
Точка 2. Вихідний спектр



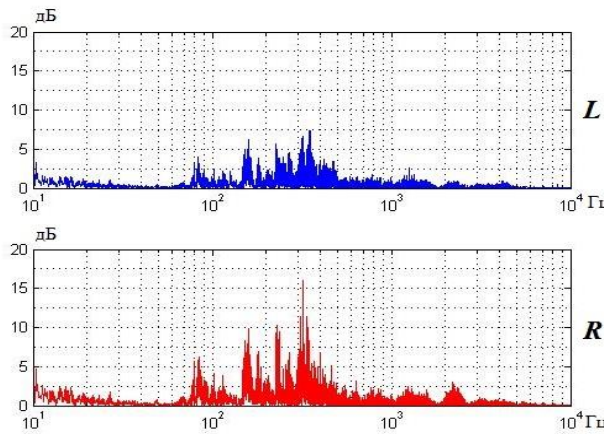
Точка 1. Вихідний спектр



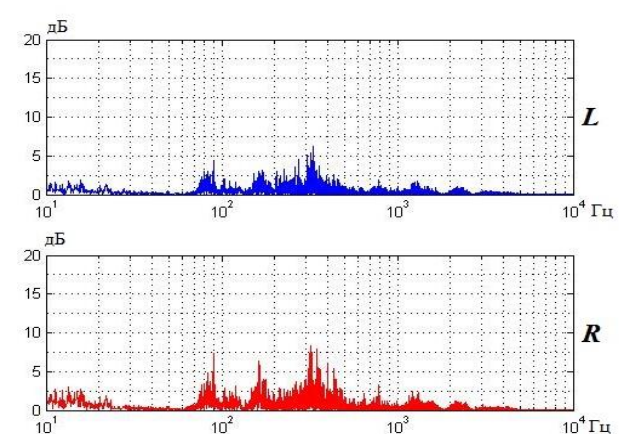
Точка 4. Вихідний спектр



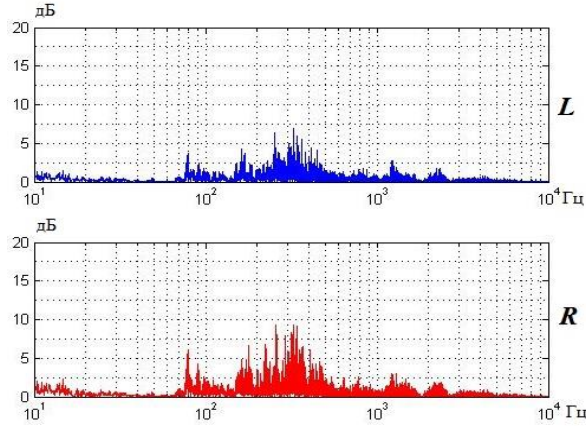
Точка 5. Вихідний спектр



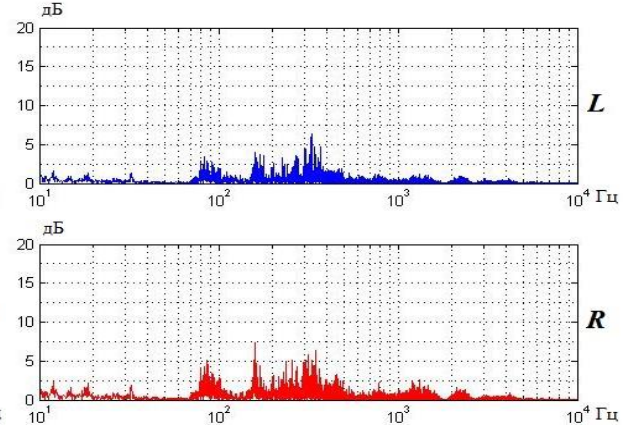
Точка 8. Вихідний спектр



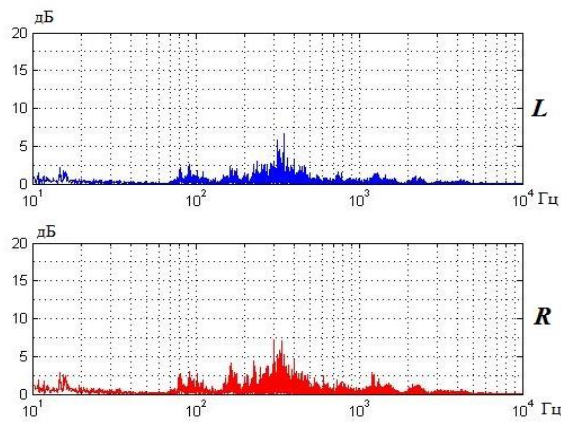
Точка 6. Вихідний спектр



Точка 9. Вихідний спектр



Точка 10. Вихідний спектр



Точка 12. Вихідний спектр

